Agenzia per l'Italia Digitale

Linee Guida recanti regole tecniche per l'apertura dei dati e il riutilizzo dell'informazione del settore pubblico

Art. 12 D.Lgs. n. 36/2006 e s.m.i.

versione 1.0

Sommario

Gruppo di lavoro

La redazione del documento è stata curata dal Gruppo di lavoro coordinato da AgID e composto dai rappresentanti dei seguenti Enti:

- **Agenzia per l'Italia Digitale;**
- **Dipartimento per la Trasformazione Digitale della Presidenza del Consiglio dei Ministri**
- **Formez PA;**
- **Istituto Geografico Militare;**
- **Università degli Studi di Roma La Sapienza;**
- **Università degli Studi di Roma Tre;**
- **CSI Piemonte;**
- **Regione Basilicata;**
- **Regione Calabria;**
- **Regione del Veneto;**
- **Regione Emilia-Romagna;**
- **Regione Friuli-Venezia Giulia;**
- **Regione Lazio;**
- **Regione Lombardia;**
- **Regione Marche;**
- **Regione Molise;**
- **Regione Piemonte;**
- **Regione Puglia;**
- **Regione Toscana;**
- **Provincia Autonoma di Trento.**

Ha collaborato, inoltre, su specifici aspetti, il **Ministero della Cultura**.

Prefazione

Le presenti Linee Guida rappresentano l'attuazione dell'art. 12 del D. Lgs. 36/2006, come modificato dal D. Lgs. 200/2021, recepimento della Direttiva (UE) 2019/1024 (c.d. Direttiva Open Data).

Esse vengono emesse ai sensi dell'articolo 71 del D. Lgs. 82/2005 e s.m.i. recante "Codice dell'Amministrazione Digitale" (di seguito **CAD**) e della Determinazione AgID n. 160/2018 recante "*Regolamento per l'adozione di linee guida per l'attuazione del Codice dell'Amministrazione Digitale*"[1]. Ai sensi del citato art. 71, divengono efficaci il giorno successivo alla loro pubblicazione sul sito istituzionale di AgID e di esse è data notizia sulla Gazzetta Ufficiale della Repubblica italiana.

Le presenti Linee guida sostituiscono, altresì, le *Linee Guida per la valorizzazione del patrimonio informativo pubblico*, adottate con Determinazione Commissariale di AgID n. 95/2014[2].

[1] https://trasparenza.agid.gov.it/archivio19_regolamenti_0_5376.html

[2] https://www.agid.gov.it/sites/default/files/repository_files/leggi_decreti_direttive/dt_cs_95_2014dig_linee_guida_open_data_art_52.pdf

Introduzione

La costruzione di un mercato unico dei dati e di un'economia basata sugli stessi è l'obiettivo strategico che l'Unione Europea intende perseguire per agevolare la circolazione dei dati a beneficio di tutti, imprese, ricercatori, amministrazioni pubbliche e cittadini. Molteplici sono le iniziative che l'UE sta portando avanti in tale direzione, fra cui la Strategia Europea per i dati[3] e la costruzione di spazi di dati comuni e interoperabili, il Regolamento relativo alla governance dei dati (Data Governance Act)[4], avente ad oggetto la facilitazione della condivisione tra settori e Stati Membri, e la proposta relativa al Data Act[5], finalizzato a introdurre norme armonizzate per l'accesso equo ai dati e il loro utilizzo.

In questo contesto si pone anche la Direttiva (UE) 2019/1024, cosiddetta Direttiva Open Data, relativa all'apertura dei dati e al riutilizzo dell'informazione del settore pubblico, con cui viene aggiornata e modificata la normativa previgente, definita a livello europeo sin dal 2003 con la prima Direttiva PSI. Le modifiche si sono rese necessarie a seguito della crescita esponenziale della quantità di dati, compresi i dati pubblici, della produzione e raccolta di nuovi tipi di dati, insieme a un'evoluzione costante delle tecnologie per l'analisi, lo sfruttamento e l'elaborazione dei dati, quali l'apprendimento automatico, l'intelligenza artificiale e l'internet delle cose (cfr. Considerando 10 della Direttiva).

In particolare, tali modifiche hanno riguardato i seguenti aspetti: la disponibilità in tempo reale dei dati dinamici mediante mezzi tecnici adeguati; l'aumento dell'offerta di dati pubblici preziosi a fini di riutilizzo, anche provenienti da imprese pubbliche, organizzazioni che svolgono attività di ricerca e organizzazioni che finanziano la ricerca; il contrasto dell'emergere di nuove forme di accordi di esclusiva; il ricorso a deroghe al principio della tariffazione dei costi marginali; la relazione tra la direttiva e alcuni strumenti giuridici correlati, tra cui la Direttiva 2007/2/CE, cosiddetta Direttiva INSPIRE (cfr. Considerando 4 della Direttiva).

Anche in Italia, come indicato anche nel Piano Triennale dell'informatica nella Pubblica Amministrazione[6], la valorizzazione del patrimonio informativo pubblico è un obiettivo strategico per la Pubblica Amministrazione per affrontare efficacemente le nuove sfide derivanti dalle

[3] https://eur-lex.europa.eu/legal-content/IT/TXT/PDF/?uri=CELEX:52020DC0066&from=IT
[4] https://eur-lex.europa.eu/legal-content/IT/TXT/PDF/?uri=CELEX:32022R0868&from=EN
[5] https://eur-lex.europa.eu/legal-content/IT/TXT/PDF/?uri=CELEX:52022PC0068&from=IT
[6] https://www.agid.gov.it/sites/default/files/repository_files/piano_triennale_per_linformatica_nella_pubblica_amministrazione_2021-2023.pdf

strategie europee innanzi delineate, garantire la creazione di servizi digitali a valore aggiunto per cittadini, imprese e, in generale, tutti i portatori di interesse e fornire ai policy maker strumenti *data-driven* da utilizzare nei processi decisionali e/o produttivi. Questo obiettivo si può perseguire solo se viene garantito l'accesso a un sempre maggiore numero di dati e vengono migliorate le modalità e la capacità di utilizzarli.

Va in questa direzione il recepimento della Direttiva (UE) 2019/1024, avvenuto con il D. Lgs. 200/2021, che ha modificato, in alcune parti anche sostanzialmente, il D. Lgs. 36/2006 che diventa, quindi, il riferimento normativo nazionale in tema di apertura di dati e riutilizzo del patrimonio informativo pubblico.

L'art. 12 del D. Lgs. 36/2006 dispone che l'Agenzia per l'Italia Digitale adotti "*le Linee guida contenenti le regole tecniche per l'attuazione del presente decreto con le modalità previste dall'articolo 71 del decreto legislativo 7 marzo 2005, n. 82*". Il presente documento rappresenta, pertanto, l'attuazione di detto articolo ed è finalizzato a definire le regole tecniche per l'implementazione delle previsioni normative introdotte dal decreto di cui sopra.

In quanto documento attuativo e considerato il carattere vincolante delle disposizioni in esso contenute anche alla luce del secondo periodo del citato art. 12 (a norma del quale "*In caso di violazione delle disposizioni introdotte dalle Linee guida, il soggetto interessato può rivolgersi al difensore civico per il digitale di cui all'articolo 17, comma 1-quater, del Codice dell'amministrazione digitale e si applicano le sanzioni previste dall'articolo 18-bis, comma 5, dello stesso Codice*"), l'approccio seguito per la redazione è stato quello di definire opportuni requisiti, quindi indicazioni da implementare obbligatoriamente, solo relativamente agli aspetti coperti dal decreto. In particolare, quindi:

- formati e modalità di pubblicazione dei dati di tipo aperto, in generale, e delle specifiche categorie di dati individuate, in particolare;
- richieste di riutilizzo;
- licenze, tariffazione e accordi di esclusiva;
- strumenti di ricerca.

Altri aspetti comunque ritenuti importanti per il processo di apertura dei dati sono stati trattati come raccomandazioni.

Capitolo 1

Ambito di applicazione

Le presenti Linee Guida stabiliscono le regole tecniche per l'attuazione del D. Lgs. 24 gennaio 2006, n. 36 (di seguito **Decreto**), così come modificato dal D. Lgs. 8 novembre 2021, n. 200 di recepimento della Direttiva (UE) 2019/1024 (di seguito **Direttiva**), in ottemperanza all'art. 12 del Decreto stesso.

Il documento ha l'obiettivo di supportare le pubbliche amministrazioni e gli altri soggetti interessati (v. par. 1.3) nel processo di apertura dei dati e di riutilizzo dell'informazione del settore pubblico attraverso indicazioni volte ad implementare le disposizioni e le modalità disciplinate dal Decreto.

In particolare, tali indicazioni riguardano formati, modalità di pubblicazione, profili di metadati, licenze e tariffazione, richieste di riutilizzo e strumenti di ricerca, tutti aspetti regolamentati da Direttiva e Decreto. Il presente documento include, inoltre, raccomandazioni su aspetti organizzativi e qualità dei dati.

1.1 Documenti oggetto di applicazione

Coerentemente con quanto disposto dall'art. 1 del Decreto, **le presenti Linee Guida si applicano**:

- a tutti i dati pubblici e tutti i documenti contenenti dati pubblici detenuti da pubbliche amministrazioni, organismi di diritto pubblico, imprese pubbliche e private, ad eccezione di quelli indicati nel par. 1.2;
- a dati e documenti i cui diritti di proprietà intellettuale sono detenuti da biblioteche, comprese le biblioteche universitarie, i musei e gli archivi, solo se il riutilizzo è autorizzato in conformità a quanto previsto nella Parte II, Titolo II, Capo I e Capo III, del decreto legislativo 22 gennaio 2004, n. 42 e nella Parte II, Titolo VII, Capo II, del decreto legislativo 30 giugno 2003 n. 196;
- ai dati della ricerca (v. par. 4.4);
- ai dati territoriali ai quali si applica il D Lgs. 27 gennaio 2010, n. 32, di recepimento della Direttiva 2007/2/CE, cosiddetta Direttiva INSPIRE (v. par. 4.5);

- ai documenti che all'atto di entrata in vigore delle indicazioni del Decreto sono già stati diffusi per il riutilizzo.

Quando, nel prosieguo del presente documento, si utilizzeranno i termini "documenti" e "dati" si intenderà fare riferimento a quei documenti e dati che rientrano nell'ambito di applicazione delineato.

1.2 Documenti esclusi dall'applicazione

Coerentemente con quanto disposto dall'art. 3 del Decreto, **le presenti Linee Guida non si applicano a dati e documenti**:

- detenuti per finalità che esulano dall'ambito dei compiti istituzionali della pubblica amministrazione o dell'organismo di diritto pubblico;
- nella disponibilità di imprese pubbliche, prodotti al di fuori dell'ambito della prestazione di servizi di interesse generale;
- nella disponibilità di imprese pubbliche, connessi ad attività direttamente esposte alla concorrenza e non soggette alle norme in materia di appalti;
- nella disponibilità delle emittenti di servizio pubblico e delle società da esse controllate e da altri organismi o loro società controllate per l'adempimento di un compito di radiodiffusione di servizio pubblico;
- nella disponibilità di istituti di istruzione secondaria e inferiore;
- diversi da quelli della ricerca, nel caso di tutti gli altri istituti di istruzione;
- nella disponibilità di enti culturali diversi dalle biblioteche, comprese quelle universitarie, dai musei e dagli archivi[7];
- nella disponibilità degli organismi coinvolti nel sistema di informazione per la sicurezza della Repubblica;
- esclusi, ai sensi dell'articolo 24 della legge 7 agosto 1990, n. 241, anche solo per un limitato periodo di tempo fissato dalle amministrazioni, ovvero:
    - documenti coperti da segreto di Stato ai sensi della legge 24 ottobre 1977, n. 801 e nei casi di segreto o di divieto di divulgazione espressamente previsti dalla legge,

[7] Il Considerando (65) della Direttiva indica alcuni esempi di enti culturali che devono essere esclusi dall'applicazione della Direttiva stessa, ovvero orchestre, teatri lirici, compagnie di ballo e teatri, compresi gli archivi che ne fanno parte, in virtù della loro specificità di "arti dello spettacolo" e del fatto che quasi tutto il loro materiale è soggetto a diritti di proprietà intellettuale di terzi.

dal regolamento governativo adottato ai sensi dell'articolo 17, comma 2, della legge 23 agosto 1988, n. 400 e dalle pubbliche amministrazioni;
  - documenti dei procedimenti tributari;
  - documenti delle attività di formazione ed emanazione di atti normativi, amministrativi generali, di pianificazione e di programmazione;
  - documenti dei procedimenti selettivi, nei confronti dei documenti amministrativi contenenti informazioni di carattere psicoattitudinale relativi a terzi;
- esclusi ai sensi dell'articolo 5-bis del decreto legislativo 14 marzo 2013, n. 33, ovvero:
  - documenti o parti di documenti inerenti ai seguenti interessi pubblici:
    - sicurezza pubblica e ordine pubblico;
    - sicurezza nazionale;
    - difesa e questioni militari;
    - relazioni internazionali;
    - politica e stabilità finanziaria ed economica dello Stato;
    - conduzione di indagini sui reati e loro perseguimento;
    - regolare svolgimento di attività ispettive;
  - documenti o parti di documenti inerenti ai seguenti interessi privati:
    - protezione dei dati personali, in conformità con la disciplina legislativa in materia;
    - libertà e segretezza della corrispondenza;
    - interessi economici e commerciali di una persona fisica o giuridica, ivi compresi la proprietà intellettuale, il diritto d'autore e i segreti commerciali;
- esclusi per motivi di tutela del segreto statistico ai sensi dell'art. 9 del decreto legislativo 6 settembre 1989, n. 322, ovvero dati raccolti nell'ambito di rilevazioni statistiche comprese nel programma statistico nazionale da parte degli uffici di statistica. Tali dati possono essere resi disponibili solo in forma aggregata, in modo che non se ne possa trarre alcun riferimento relativamente a persone identificabili, e possono essere utilizzati solo per scopi statistici;
- su cui terzi detengono diritti di proprietà intellettuale ai sensi della legge 22 aprile 1941, n. 633, ovvero diritti di proprietà industriale ai sensi del decreto legislativo 10 febbraio 2005, n. 30;
- non contenenti dati pubblici, il cui accesso è disciplinato dal Capo V della legge 7 agosto 1990, n. 241;

- parti di documenti contenenti solo logotipi, stemmi e distintivi;
- il cui accesso, ai sensi delle previsioni del Regolamento (UE) 2016/679 e del decreto legislativo 30 giugno 2003, n. 196 e s.m.i. nonché del decreto legislativo 18 maggio 2018, n. 51, è escluso, limitato o risulti pregiudizievole per la tutela della vita privata e dell'integrità degli individui, comprese le parti di documenti accessibili che contengono dati personali il cui riutilizzo è stato definito per legge incompatibile con le previsioni delle suddette disposizioni normative;
- il cui accesso è escluso o limitato per motivi di protezione delle informazioni sensibili relative alle infrastrutture critiche definite, all'articolo 2 del decreto legislativo 11 aprile 2011, n. 61, come "*infrastrutture, ubicate in uno Stato membro dell'Unione europea, che sono essenziali per il mantenimento delle funzioni vitali della società, della salute, della sicurezza e del benessere economico e sociale della popolazione ed il cui danneggiamento o la cui distruzione avrebbe un impatto significativo in quello Stato, a causa dell'impossibilità di mantenere tali funzioni";*
- diversi dai dati della ricerca, nella disponibilità di organizzazioni che svolgono attività di ricerca e di organizzazioni che finanziano la ricerca, comprese le organizzazioni preposte al trasferimento dei risultati della ricerca.

1.3 Soggetti destinatari

I destinatari delle presenti Linee Guida sono quelli individuati dal Decreto e innanzi già citati, ovvero le pubbliche amministrazioni, gli organismi di diritto pubblico e le imprese pubbliche e private.

Per l'individuazione delle "**pubbliche amministrazioni**" si fa riferimento all'art. 1, comma 2 del D. Lgs. 30 marzo 2001, n. 165; sono compresi, quindi:

- tutte le amministrazioni dello Stato, ivi compresi gli istituti e scuole di ogni ordine e grado e le istituzioni educative;
- le aziende ed amministrazioni dello Stato ad ordinamento autonomo;
- le Regioni;
- le Province e le Città Metropolitane;
- i Comuni, le Comunità montane, e loro consorzi e associazioni;
- le istituzioni universitarie;
- gli Istituti autonomi case popolari;
- le Camere di commercio, industria, artigianato e agricoltura e loro associazioni;

- tutti gli enti pubblici non economici nazionali, regionali e locali;
- le amministrazioni, le aziende e gli enti del Servizio sanitario nazionale;
- l'Agenzia per la rappresentanza negoziale delle pubbliche amministrazioni (ARAN);
- le Agenzie di cui al decreto legislativo 30 luglio 1999, n. 300.

Il Decreto precisa che sono comprese anche le autorità di sistema portuale, le autorità amministrative indipendenti di garanzia, vigilanza e regolazione, nonché i loro consorzi o associazioni a qualsiasi fine istituiti.

Per l'individuazione degli "**organismi di diritto pubblico**" si fa riferimento all'art. 3, comma 1, lettera d) del D. Lgs. 18 aprile 2016, n. 50; sono, quindi, compresi, sulla base dell'allegato IV del predetto decreto legislativo, i seguenti organismi[8]:

- Mostra d'oltremare S.p.A.;
- Ente nazionale per l'aviazione civile - ENAC;
- Società nazionale per l'assistenza al volo S.p.A. - ENAV;
- ANAS S.p.A.;
- Consip S.p.A. (quando Consip agisce in qualità di centrale di committenza per le autorità sub-centrali)

e le seguenti categorie:

- Consorzi per le opere idrauliche;
- Università statali, gli istituti universitari statali, i consorzi per i lavori interessanti le università;
- Istituzioni pubbliche di assistenza e di beneficenza;
- Istituti superiori scientifici e culturali, osservatori astronomici, astrofisici, geofisici o vulcanologici;
- Enti di ricerca e sperimentazione;
- Enti che gestiscono forme obbligatorie di previdenza e di assistenza;
- Consorzi di bonifica;
- Enti di sviluppo e di irrigazione;
- Consorzi per le aree industriali;
- Comunità montane;
- Enti preposti a servizi di pubblico interesse;

[8] Alcuni organismi o categorie possono corrispondere a pubbliche amministrazioni indicate innanzi.

- Enti pubblici preposti ad attività di spettacolo, sportive, turistiche e del tempo libero;
- Enti culturali e di promozione artistica.

Per quanto riguarda le **imprese pubbliche**, sono da considerare:

- quelle attive nei settori di cui agli articoli da 115 a 121 del D. Lgs. 18 aprile 2016, n. 50, che, all'atto di adottare le presenti Linee Guida, sono i seguenti:
  - gas ed energia termica;
  - elettricità;
  - acqua;
  - servizi di trasporto;
  - porti e aeroporti;
  - servizi postali;
  - estrazione di gas e prospezione o estrazione di carbone o di altri combustibili solidi;
- quelle che agiscono in qualità di *"operatore di servizio pubblico"*, definito dall'art. 2 del Regolamento (CE) n. 1370/2007 come "*un'impresa o un gruppo di imprese di diritto pubblico o privato che fornisce servizi di trasporto pubblico di passeggeri o qualsiasi ente pubblico che presta servizi di trasporto pubblico di passeggeri*";
- quelle che agiscono in qualità di vettori aerei che assolvono oneri di servizio pubblico ai sensi dell'art. 16 del Regolamento (CE) n. 1008/2008;
- quelle che agiscono in qualità di armatori comunitari che assolvono obblighi di servizio pubblico ai sensi dell'art. 4 del Regolamento (CEE) n. 3577/1992.

Per quanto riguarda le **imprese private**, sono da considerare:

- le imprese private di trasporto che sono soggette ad obblighi di servizio pubblico ai sensi dell'art. 16 del Regolamento (CE) n. 1008/2008 relativo alla prestazione di servizi aerei nella Comunità Europea;
- in generale, i gestori di servizi pubblici in relazione ai servizi di pubblico interesse.

Risorse utili – cap. 1

- *Linee guida in materia di trattamento di dati personali contenuti anche in atti e documenti amministrativi effettuato da soggetti pubblici per finalità di pubblicazione e diffusione sul web*, adottate con Deliberazione n. 243 del 15 maggio 2014 del Garante per la Protezione dei Dati Personali
- *Linee Guida recanti indicazioni operative ai fini della definizione delle esclusioni e dei limiti all'accesso civico di cui all'art. 5 co. 2 del D. Lgs. 33/2013* adottate con Determinazione n. 1309 del 28 dicembre 2016 dell'Autorità Nazionale Anticorruzione
- *Piano nazionale di digitalizzazione del patrimonio culturale – Linee Guida per l'acquisizione, la circolazione e il riuso delle riproduzioni dei beni culturali in ambiente digitale*, Ministero della Cultura

Capitolo 2

Normativa di riferimento

Sono riportati di seguito gli atti normativi di riferimento del presente documento.

2.1 Legislazione nazionale ed europea

[CAD] Decreto legislativo 7 marzo 2005, n. 82 recante "*Codice dell'amministrazione digitale*"

[D-LGS-196-2003] Decreto legislativo 30 giugno 2003, n. 196 e s.m.i. recante "*Codice in materia di protezione dei dati personali, recante disposizioni per l'adeguamento dell'ordinamento nazionale al regolamento (UE) n. 2016/679 del Parlamento europeo e del Consiglio, del 27 aprile 2016, relativo alla protezione delle persone fisiche con riguardo al trattamento dei dati personali, nonché alla libera circolazione di tali dati e che abroga la direttiva 95/46/CE*"

[D-LGS-36-2006] Decreto legislativo 24 gennaio 2006, n. 36 recante "*Attuazione della direttiva (UE) 2019/1024 relativa all'apertura dei dati e al riutilizzo dell'informazione del settore pubblico che ha abrogato la direttiva 2003/98/CE*"

[D-LGS-32-2010] Decreto legislativo 27 gennaio 2010, n. 32 recante "*Attuazione della direttiva 2007/2/CE, che istituisce un'infrastruttura per l'informazione territoriale nella Comunità europea (INSPIRE)*"

[D-LGS-33-2013] Decreto legislativo 14 marzo 2013, n. 33 recante "*Riordino della disciplina riguardante il diritto di accesso civico e gli obblighi di pubblicità, trasparenza e diffusione di informazioni da parte delle pubbliche amministrazioni*"

[D-LGS-51-2018] Decreto legislativo 18 maggio 2018, n. 51 recante "*Attuazione della direttiva (UE) 2016/680 del Parlamento europeo e del Consiglio, del 27 aprile 2016, relativa alla protezione delle persone fisiche con riguardo al trattamento dei dati personali da parte delle autorità competenti a fini di prevenzione, indagine, accertamento e perseguimento di reati o esecuzione di sanzioni penali, nonché alla libera circolazione di tali dati e che abroga la decisione quadro 2008/977/GAI del Consiglio*"

[D-LGS-200-2021] Decreto legislativo 8 novembre 2021, n. 200 recante "*Attuazione della direttiva (UE) 2019/1024 del Parlamento europeo e del Consiglio, del 20 giugno 2019, relativa all'apertura dei dati e al riutilizzo dell'informazione del settore pubblico (rifusione)*"

[INSPIRE-DIR] Direttiva 2007/2/CE del Parlamento europeo e del Consiglio, del 14 marzo 2007, che istituisce un'Infrastruttura per l'informazione territoriale nella Comunità europea (INSPIRE)

[DIR-2016-680] Direttiva (UE) 2016/680 del Parlamento europeo e del Consiglio del 27 aprile 2016 relativa alla protezione delle persone fisiche con riguardo al trattamento dei dati personali da parte delle autorità competenti a fini di prevenzione, indagine, accertamento e perseguimento di reati o esecuzione di sanzioni penali, nonché alla libera circolazione di tali dati e che abroga la decisione quadro 2008/977/GAI del Consiglio

[OD-DIR] Direttiva (UE) 2019/1024 del Parlamento Europeo e del Consiglio del 20 giugno 2019 relativa all'apertura dei dati e al riutilizzo dell'informazione del settore pubblico (rifusione)

[GDPR] Regolamento (UE) 2016/679 del Parlamento europeo e del Consiglio del 27 aprile 2016 relativo alla protezione delle persone fisiche con riguardo al trattamento dei dati personali, nonché alla libera circolazione di tali dati e che abroga la direttiva 95/46/CE (regolamento generale sulla protezione dei dati)

[REG-868-22] Regolamento (UE) 2022/868 del Parlamento europeo e del Consiglio del 30 maggio 2022 relativo alla governance europea dei dati e che modifica il regolamento (UE) 2018/1724 (Regolamento sulla governance dei dati)

[REG-HVD] Regolamento di esecuzione (UE) 2023/138 della Commissione del 21 dicembre 2022 che stabilisce un elenco di specifiche serie di dati di elevato valore e le relative modalità di pubblicazione e riutilizzo

2.2 Linee Guida

[LG-CONS] *Linee Guida sulla formazione, gestione e conservazione dei documenti informatici,* adottate con la Determinazione AgID n. 407/2020 come modificata con la Determinazione AgID n. 371/2021

[LG-INT] *Linee Guida sull'interoperabilità tecnica delle Pubbliche Amministrazioni,* adottate con la Determinazione AgID n. 547/2021 del 1° ottobre 2021

[LG-PDND] *Linee Guida sull'infrastruttura tecnologica della Piattaforma Digitale Nazionale Dati per l'interoperabilità dei sistemi informativi e delle basi di dati,* adottate con la Determinazione AgID n. 627/2021 del 15 dicembre 2021

[LG-RNDT] *Linee Guida recanti regole tecniche per la definizione e l'aggiornamento del contenuto del Repertorio Nazionale dei Dati Territoriali,* adottate con la Determinazione AgID n. 50/2022 del 28 febbraio 2022

[LG-SIC] *Linee Guida Tecnologie e standard per la sicurezza dell'interoperabilità tramite API dei sistemi informatici,* adottate con la Determinazione AgID n. 547/2021 del 1° ottobre 2021

[LG-GPDP] *Linee guida in materia di trattamento di dati personali, contenuti anche in atti e documenti amministrativi, effettuato per finalità di pubblicità e trasparenza sul web da soggetti pubblici e da altri enti obbligati*, adottate con provvedimento del Garante per la protezione dei dati personali n. 243 del 15 maggio 2014

Si specifica che tali ultime Linee guida sono attualmente in corso di aggiornamento con specifici riferimenti anche al tema del riutilizzo.

Si richiamano altresì:

[PAR-06-2013] Parere 06/2013 del Gruppo di lavoro Articolo 29 sul riutilizzo dei dati aperti e delle informazioni del settore pubblico

[PAR-05-2014] Parere 05/2014 del Gruppo di lavoro Articolo 29 sulle tecniche di anonimizzazione

[PAR-03-2021] Parere congiunto del Comitato europeo per la protezione dei dati e del Garante europeo della protezione dei dati 03/2021 sulla proposta di regolamento del Parlamento europeo e del Consiglio relativo alla governance europea dei dati (Atto sulla governance dei dati)

Capitolo 3

Terminologia

3.1 Note di lettura del documento

Conformemente alle norme *ISO/IEC Directives, Part 3* per la stesura dei documenti tecnici le presenti Linee Guida utilizzeranno le parole chiave «DEVE», «DEVONO», «NON DEVE», «NON DEVONO», «È RICHIESTO», «DOVREBBE», «NON DOVREBBE», «È RACCOMANDATO», «NON È RACCOMANDATO» «PUÒ» e «È OPZIONALE», la cui interpretazione è descritta di seguito.

- **DEVE** o **DEVONO** o **È RICHIESTO**, indicano un requisito obbligatorio per rispettare le Linee Guida;
- **NON DEVE** o **NON DEVONO**, indicano un assoluto divieto delle specifiche;
- **DOVREBBE** o **È RACCOMANDATO/SI RACCOMANDA** o **NON DOVREBBE** o **NON È RACCOMANDATO**, indicano che le implicazioni devono essere comprese e attentamente pesate prima di scegliere approcci alternativi;
- **PUÒ** o **POSSONO** o **È OPZIONALE**, indica che si può scegliere di applicare o meno senza alcun tipo di implicazione la specifica.

Ove ricorrano le condizioni, nei prossimi capitoli sono definiti specifici requisiti e raccomandazioni in modo da rendere in maniera più chiara e sintetica **cosa è necessario implementare obbligatoriamente** per rispettare le norme giuridiche richiamate dalle presenti Linee Guida (requisiti) e **cosa invece è solo un forte suggerimento** (raccomandazioni). Nel caso dei requisiti si utilizzeranno le parole chiave indicate innanzi ai primi due punti, mentre nel caso delle raccomandazioni si utilizzeranno quelle indicate agli ultimi due punti.

I requisiti, inoltre, saranno rappresentati nel modo seguente:

REQUISITO x: dlgs36-2006/opendata/req/<conformance-class-id>/<req-id>
Testo del requisito […]

mentre le raccomandazioni nel modo seguente:

Raccomandazione x: dlgs36-2006/opendata/rec/<conformance-class-id>/<rec-id>
Testo della raccomandazione […].

Quando, su alcuni aspetti specifici, esistono già documenti (norme, regole tecniche, linee guida, manuali, materiali vari) che approfondiscono tali aspetti e che forniscono indicazioni puntuali, anche eventualmente vincolanti, è presente anche la **sezione "Risorse utili"** che riporta i riferimenti, attraverso collegamenti ipertestuali, a tali documenti o risorse in generale a cui si rimanda. In particolare, i documenti e le risorse possono essere relativi ad una di queste tipologie:

- linee guida, regolamenti o altre tipologie di documenti che disciplinano aspetti specifici trattati nel presente documento e che non si ritiene di dover duplicare (per evitare, per esempio, disallineamenti in caso di aggiornamento degli stessi), anche perché in alcuni casi di competenza di specifiche amministrazioni;
- guide, specifiche, documenti tecnici, manuali e altri materiali formativi che forniscono indicazioni tecniche puntuali, specie se definiti nell'ambito del framework europeo che non si ritiene di scalare o ridefinire a livello nazionale per garantire la coerenza e l'interoperabilità in un contesto superiore.

Tale sezione è rappresentata come segue:

Risorse utili – par. x.x

- link a risorsa 1 …
- link a risorsa 2 …
- …

3.2 Termini e definizioni

Ai fini del presente documento, oltre alle definizioni pertinenti di cui all'art. 2 del Decreto e all'art. 1 del **CAD**, si applicano le seguenti definizioni:

3.2.1
catalogo dati

strumento di gestione dei metadati utilizzato per repertoriare e organizzare i dati all'interno dei sistemi di un'organizzazione

3.2.2
collective database

database in forma non modificata come parte di una raccolta di database indipendenti che insieme sono assemblati in una unica collezione. Un'opera che costituisce un "collective database" (Banca Dati Collettiva) non sarà considerata una Banca Dati Derivata [ODbL license]

3.2.3

dataset

collezione identificabile di dati [ISO]

collezione di dati, pubblicati o curati da un'unica fonte, e disponibili per l'accesso o il download in uno o più formati [DCAT-AP]

3.2.4

download in blocco

operazione che consente, in modo semplice ed efficiente, di scaricare, con un'unica richiesta, l'intero set di dati nel sistema dell'utente

3.2.5

DRM clause (clausola DRM)

"*eventuali termini o misure tecnologiche presenti nel Database, in un Database Derivato, o in tutto o in parte sostanziale dei contenuti che alterano o limitano i termini della Licenza*" [licenza ODbL], intesa, quindi, come divieto di imposizione salvo possibilità di lasciare copia aperta

"*misure che, in assenza di apposita autorità, possono non essere eluse ai sensi di leggi che adempiono agli obblighi di cui all'articolo 11 del Trattato WIPO sul diritto d'autore adottato il 20 dicembre 1996, e/o accordi internazionali simili*" [licenza CC-BY], intesa quindi come divieto di imposizione senza alternative

3.2.6

portale dati

un sistema basato sul Web che contiene un catalogo dati con le descrizioni dei dataset e fornisce servizi che consentono la ricerca e il riutilizzo dei dataset stessi [DCAT-AP]

3.2.7

processo di apertura dei dati

processo volto alla produzione e alla disponibilità di dati basato sul principio dell'"*apertura fin dalla progettazione e per impostazione predefinita*" dell'informazione del settore pubblico, in formati aperti che possono essere utilizzati, riutilizzati e condivisi liberamente da chiunque e per qualsiasi finalità, nei limiti previsti dalla disciplina vigente. L'apertura dei dati, dunque, riguarda la "messa a disposizione" dei dati del settore pubblico per il riutilizzo, operazione che non coincide

necessariamente con la "pubblicazione" dei dati, ma anche con la relativa fruizione conseguente alla "comunicazione" degli stessi a seguito di specifica richiesta di riutilizzo

3.2.8

produced work

un'opera (come un'immagine, materiale audiovisivo, testo o suoni) risultante dall'utilizzo della totalità o di una parte sostanziale dei contenuti (tramite una ricerca o altra query) da un Database, inteso come un Database Derivato, o da un Database inteso come parte di una banca dati collettiva (collective database) [ODbL license]

3.2.9

raccomandazione

possibile scelta suggerita senza necessariamente escluderne altre

3.2.10

requisito

criterio da soddisfare per garantire la conformità al documento e da cui non è consentita alcuna deviazione

3.2.11

result from computational use

risultati o output che l'utente ottiene dall'analisi (attraverso l'uso di dispositivi di calcolo o altro) o da altre interpretazioni dei dati

3.2.12

riutilizzatore

una persona o un'organizzazione che utilizza i dati (aperti) disponibili per i propri scopi

3.3 Acronimi

Di seguito si riportano gli ACRONIMI che sono utilizzati nelle presenti Linee Guida.

API	Application Programming Interface
CC	Creative Commons
CMS	Content Management System
CSV	Comma Separated Value

DCAT	Data Catalog Vocabulary
DCAT-AP	Data Catalog Vocabulary - Application Profile
DCAT-AP_IT	Data Catalog Vocabulary - Application Profile ITaliano
DRM	Digital Rights Management
FAIR	Findable Accessible Interoperable Reusable
http	HyperText Transfer Protocol
INSPIRE	INfrastructure for SPatial InfoRmation in Europe
ICT	Information and Communication Technology
IGM	Istituto Geografico Militare
IPA	Indice dei domicili digitali della Pubblica Amministrazione e dei Gestori di Pubblici Servizi
ISA	Interoperability Solutions for public Administration
LOD	Linked Open Data
JSON	JavaScript Object Notation
OD	Open Data
OWL	Ontology Web Language
PA	Pubblica Amministrazione
PDND	Piattaforma Digitale Nazionale Dati
PMI	Piccole e Medie Imprese
PSI	Public Sector Information
RDF	Resource Description Framework
RDFS	RDF Schema
RNDT	Repertorio Nazionale Dati Territoriali
RTD	Responsabile per la Transizione Digitale

SDMX Statistical Data and Metadata eXchange

SPARQL Sparql Protocol And Rdf Query Language

URI Uniform Resource Identifier

XML eXtensible Markup Language

Capitolo 4

Principi generali

L'art. 6 del Decreto stabilisce che pubbliche amministrazioni, organismi di diritto pubblico, imprese pubbliche e private, che rientrano tra i soggetti destinatari di cui al par. **1.3**, rendano disponibili per il riutilizzo i propri dati, e relativi metadati, nel rispetto delle regole tecniche definite con le presenti Linee Guida.

Questo capitolo definisce le modalità e i formati per i dati da rendere disponibili come dati di tipo aperto e, in particolare, per le specifiche categorie di dati individuate da Direttiva e Decreto, ovvero dati dinamici, dati della ricerca e serie di dati di elevato valore (di cui, rispettivamente, agli artt. 6, commi 5, 6, 9-bis e 12-bis del Decreto).

4.1 Requisiti comuni

Sulla base dell'art. 1 del Decreto, le presenti Linee Guida si applicano ai "*documenti contenenti* ***dati pubblici*** *nella* ***disponibilità*** *delle* ***pubbliche amministrazioni****, degli* ***organismi di diritto pubblico*** *e delle* ***imprese pubbliche e private***", con esclusione dei casi riportati al par. **1.2**.

Il Decreto utilizza il termine "documento" nell'accezione di cui all'art. 2, comma 1, lettera c), cioè "la *rappresentazione di atti, fatti e dati a prescindere dal supporto, cartaceo o elettronico, registrazione sonora, visiva o audiovisiva o qualsiasi parte di tale contenuto nella disponibilità della pubblica amministrazione o dell'organismo di diritto pubblico. […]*". La definizione di documento non comprende i programmi informatici ma include qualsiasi parte del documento stesso.

Il Decreto - a differenza della Direttiva - include anche la definizione di "dato pubblico" nel medesimo articolo 2, comma 1, lettera d), quale "*dato conoscibile da chiunque*".

Non è, quindi, definita chiaramente la distinzione tra "documento" e "dato"; anzi, se si considerano altre fonti normative, come il Regolamento (UE) 2022/868 (Data Governance Act)[9] o la proposta di Data Act[10], la definizione di "dato" presente in quelle fonti è coincidente con quella di "documento" innanzi riportata. L'unica differenza è che nel caso di "dato" è considerata solo la rappresentazione digitale, mentre nel caso di "documento" è compresa anche la rappresentazione attraverso supporto cartaceo.

[9] v. https://eur-lex.europa.eu/legal-content/IT/TXT/PDF/?uri=CELEX:32022R0868&from=EN
[10] v. https://eur-lex.europa.eu/legal-content/IT/TXT/PDF/?uri=CELEX:52022PC0068&from=EN

Nella Direttiva, inoltre, i dati dinamici, i dati della ricerca e le serie di dati di elevato valore sono definiti anch'essi come "documenti".

Ai fini delle presenti Linee Guida si rende necessario fornire una distinzione in modo da rendere più agevole l'applicazione delle stesse.

Nella legislazione italiana si riscontra la presenza della definizione dei diversi tipi di documenti[11]. Il **CAD** presenta la definizione di "documento informatico" come "*il documento elettronico che **contiene** la rappresentazione informatica di atti, fatti o dati giuridicamente rilevanti*"; la legge n. 241/90 include, all'art. 22, comma 1, lettera d), la definizione di "documento amministrativo" come "*rappresentazione grafica, fotocinematografica, elettromagnetica o di qualunque altra specie del **contenuto** di atti, anche interni o non relativi ad uno specifico procedimento, detenuti da una pubblica amministrazione* […]".

Al di là dell'aggettivazione utilizzata per caratterizzare la tipologia di documento, nelle definizioni riportate dati e informazioni risultano rappresentare il contenuto di un documento e, di conseguenza, il documento risulta essere il "contenitore" di dati e informazioni. Tale interpretazione sembra essere confermata anche dall'art. 1, comma 1 del Decreto nel quale, a proposito dell'oggetto e dell'ambito di applicazione del Decreto stesso, si parla di "documenti **contenenti** dati pubblici". Nel prosieguo delle presenti Linee Guida, quindi, l'utilizzo dei termini "documento" e "dato" farà riferimento all'accezione indicata innanzi.

Considerata, quindi, la distinzione tra documenti e dati introdotta prima, l'applicazione delle indicazioni delle presenti Linee Guida non è da limitare solo ai documenti, ma DEVE riguardare anche i dati contenuti nei documenti. Ad esempio, nel processo di apertura dei dati si suggerisce, nel caso di pubblicazione di un report contenente tabelle di dati, di pubblicare separatamente sia il documento che le singole tabelle di dati ivi contenute, seguendo le regole definite nel presente documento.

A tale proposito, si precisa che quanto sopra non comporta obblighi aggiuntivi a quanto richiesto nelle "Linee Guida sulla formazione, gestione e conservazione dei documenti informatici" [**LG-CONS**] in relazione alla formazione del documento informatico. Ciò significa che le indicazioni presenti nelle citate Linee Guida continuano ad essere riferibili ai soli documenti che rientrano nell'ambito di applicazione ivi definito.

[11] v. https://www.disfipeq.unich.it/sites/st10/files/8._il_documento_informatico.pdf

In relazione, inoltre, alle modalità di formazione del documento di cui al par. 2.1.1 delle Linee Guida di cui sopra, con particolare riferimento alla lettera a)[12], nel caso in cui estrarre e pubblicare separatamente i dati contenuti nei documenti comporti difficoltà sproporzionate, è applicabile quanto indicato con il **REQUISITO 4**. SI RACCOMANDA, in questi casi, di valutare se le difficoltà sproporzionate di cui sopra possano essere superate pubblicando i documenti in modo che sia consentita l'estrazione dei dati in essi contenuti direttamente da parte dei riutilizzatori.

I dati pubblici che rientrano nell'ambito di applicazione delle presenti Linee Guida DEVONO essere messi a disposizione **per il riutilizzo a fini commerciali e non commerciali**:

- in **formato leggibile meccanicamente**, cioè, come da definizione presente nel Decreto, in "*un formato di file strutturato in modo tale da consentire alle applicazioni software di individuare, riconoscere ed estrarre facilmente dati specifici, comprese dichiarazioni individuali di fatto e la loro struttura interna*";
- in **formato aperto**, cioè, come da definizione dell'art. 1, comma 1, lettera l-bis) del **CAD**, in "*un formato di dati reso pubblico, documentato esaustivamente e neutro rispetto agli strumenti tecnologici necessari per la fruizione dei dati stessi*";
- **in modalità accessibile** attraverso le tecnologie dell'informazione e della comunicazione;
- **gratuitamente o con i costi marginali** sostenuti per la riproduzione, messa a disposizione e divulgazione dei documenti, nonché per l'anonimizzazione di dati personali o per le misure adottate per proteggere le informazioni commerciali a carattere riservato (v. par. **6.2**);
- secondo i termini di **licenze standard**, disponibili in formato digitale (v. par. **6.1**);
- provvisti dei relativi **metadati** (v. par. **4.6**).

Tali indicazioni rappresentano le caratteristiche dei **dati di tipo aperto** come da definizione fornita all'art. 1, comma 1. lettera l-ter) del **CAD**. Resta fermo che il riutilizzo dei dati pubblici DEVE essere effettuato nel rispetto della normativa unionale europea e nazionale in materia di protezione dei dati personali, con particolare attenzione all'analisi della finalità e degli eventuali limiti del riutilizzo, anche ai sensi delle presenti Linee guida.

Al riguardo, nel rispetto della normativa in materia di protezione dei dati personali, si ricorda, che i dati personali pubblicati online dalle pubbliche amministrazioni sulla base di specifici

[12] La modalità indicata è relativa alla "*creazione tramite l'utilizzo di strumenti software o servizi cloud qualificati che assicurino la produzione di documenti nei formati e nel rispetto delle regole di interoperabilità di cui all'allegato 2*"

presupposti normativi di cui all'art. 2-ter, commi 1 e 3, del D. Lgs. 196/2003 [**D-LGS-196-2003**] non sono da considerarsi automaticamente "dati di tipo aperto" liberamente riutilizzabili per qualsiasi finalità, considerando che anche la Direttiva prevede espressamente che le proprie disposizioni "*non pregiudicano il diritto dell'Unione e nazionale in materia di protezione dei dati personali, in particolare il regolamento (UE) 2016/679 e la direttiva 2002/58/CE, nonché le corrispondenti disposizioni di diritto nazionale*" (cfr. art. 1, par. 4).

Si precisa che le presenti Linee guida non introducono alcun ulteriore obbligo di pubblicazione rispetto agli obblighi già altrimenti previsti dalla normativa vigente né impongono che l'obbligo di pubblicazione online di dati personali si traduca in obbligo di pubblicazione dei medesimi dati come "dati di tipo aperto" ai sensi della definizione di cui all'art. 1, comma 1, lett. l-ter del **CAD**, anche alla luce di quanto disposto all'art. 52, comma 2 del medesimo **CAD**. In tale contesto, si evidenzia che il Decreto - facendo seguito all'art. 3, comma 1, lett. h-quater), della Direttiva - prevede che la normativa sul riutilizzo non si applichi ai documenti "*il cui accesso, ai sensi delle previsioni del regolamento UE 2016/679 e del decreto legislativo 30 giugno 2003, n. 196 nonché del decreto legislativo 18 maggio 2018, n. 51, è escluso o limitato, ovvero risulti pregiudizievole per la tutela della vita privata e dell'integrità degli individui, nonché alle parti di documenti accessibili che contengono dati personali il cui riutilizzo è stato definito per legge incompatibile con le previsioni delle suddette disposizioni normative*". Come precisato anche dal Garante per la protezione dei dati personali, la disposizione deve essere in ogni caso interpretata nel senso che il riutilizzo dei dati personali non è permesso nei casi in cui "*in particolare, non è conforme al disposto di cui agli articoli 5, paragrafo 1, lettera b) e 6, paragrafo 4, del citato Regolamento, nonché 3, comma 1, lettera b) e 6, comma 1, del decreto legislativo 18 maggio 2018, n. 51*"[13].

Il principio di limitazione della finalità (conforme al considerando n. 52 della Direttiva) DEVE supportare l'attività valutativa dell'interprete, anche rispetto all'individuazione delle più opportune azioni utili a rendere riutilizzabili i dataset, come indicato anche dal Garante per la protezione dei dati personali secondo cui "*il principio generale del libero riutilizzo di documenti contenenti dati pubblici, stabilito dalla disciplina nazionale ed europea, riguarda essenzialmente documenti che non contengono dati personali oppure riguarda dati personali opportunamente aggregati e resi anonimi*" (cfr. parte prima, par. 6, delle Linee guida [**LG-GPDP**]). Pertanto, come si legge nelle citate Linee guida del Garante, "*il semplice fatto che informazioni personali siano rese pubblicamente conoscibili online per finalità di trasparenza non comporta che le stesse siano liberamente riutilizzabili da chiunque e per qualsiasi scopo*" ma "*impone al soggetto chiamato a dare attuazione agli obblighi di pubblicazione sul proprio sito web istituzionale di determinare – qualora*

[13] Provvedimento del Garante per la protezione dei dati personali n. 308 del 26/8/2021 (v. https://www.garanteprivacy.it/web/guest/home/docweb/-/docweb-display/docweb/9717493)

intenda rendere i dati riutilizzabili – se, per quali finalità e secondo quali limiti e condizioni eventuali utilizzi ulteriori dei dati personali resi pubblici possano ritenersi leciti alla luce del 'principio di finalità' e degli altri principi di matrice europea in materia di protezione dei dati personali" (sul processo di selezione dei dati e sulle cautele da utilizzare nelle valutazioni riguardanti il riutilizzo di dati personali, si rinvia alle indicazioni contenute nel successivo par. **5.1.2)**.

Appare anche utile richiamare quanto condiviso dal Garante per la protezione dei dati personali nelle già citate Linee guida [**LG-GPDP**]: "*Pertanto, al fine di evitare di perdere il controllo sui dati personali pubblicati online in attuazione degli obblighi di trasparenza e di ridurre i rischi di loro usi indebiti, è quindi in primo luogo opportuno che le pubbliche amministrazioni e gli altri soggetti chiamati a dare attuazione agli obblighi di pubblicazione di cui al d. lgs. n. 33/2013 inseriscano nella sezione denominata "Amministrazione trasparente" dei propri siti web istituzionali un Alert generale con cui si informi il pubblico che i dati personali pubblicati sono <<riutilizzabili solo alle condizioni previste dalla normativa vigente sul riuso dei dati pubblici (direttiva comunitaria 2003/98/CE e d. lgs. 36/2006 di recepimento della stessa), in termini compatibili con gli scopi per i quali sono stati raccolti e registrati, e nel rispetto della normativa in materia di protezione dei dati personali>>*".

REQUISITO 1: dlgs36-2006/opendata/req/common/guidelines
I soggetti di cui al par. 1.3 DEVONO rendere disponibili i documenti e i dati di cui al par. 1.1 per il loro riutilizzo a fini commerciali e non commerciali secondo quanto indicato nelle presenti Linee Guida.

Nella definizione di dato di tipo aperto di cui all'art. 1, comma 1, lettera l-ter del **CAD**, è indicato esplicitamente che i dati debbano essere resi disponibili "*in formato disaggregato*". SI RACCOMANDA, quindi, di fornire dati che rispettino l'indicazione del **CAD** in tutti quei casi in cui è possibile ed è consentito dalle norme. Nel caso in cui fornire dati disaggregati implicherebbe la violazione della normativa in materia di protezione dei dati personali, è necessario adottare procedure di anonimizzazione in modo da rendere i dati non più riconducibili a una persona fisica identificata o identificabile[14]. In ogni caso, l'art. 50-ter, comma 4 del **CAD** prevede che con la definizione della Strategia nazionale dati vengano identificati le tipologie, i limiti, le finalità e le modalità di messa a disposizione, su richiesta della Presidenza del Consiglio dei ministri, dei dati aggregati e anonimizzati di cui sono titolari le Pubbliche Amministrazioni. Nel caso in cui la

[14] Si vedano sul punto il Parere 05/2014 del Gruppo di lavoro Articolo 29 sulle tecniche di anonimizzazione [**PAR-05-2014**] e il Considerando 8 del Regolamento di esecuzione (UE) 2023/138 [**REG-HVD**].

Strategia di cui sopra dovesse indicare che i dati identificati debbano essere resi disponibili come dati aperti, si applicheranno, anche in quei casi, le indicazioni fornite con le presenti Linee Guida.

Nella figura che segue è mostrata una rappresentazione delle tipologie di dati pubblici.

Il concetto di "disponibilità" può essere derivato dal **CAD** nella definizione presente prima delle modifiche apportate dal decreto legislativo 26 agosto 2016, n. 179. Sulla base della formulazione originaria del **CAD**, per disponibilità si intende "*la possibilità di accedere ai dati senza restrizioni non riconducibili a esplicite norme di legge*".

Come detto, il Decreto si applica ai **dati pubblici** nella **disponibilità** degli enti indicati al par. **1.3** e che presentino le caratteristiche dei **dati di tipo aperto** indicate innanzi (rappresentati, nella figura, dall'intersezione indicata, appunto, con "Dati di tipo aperto").
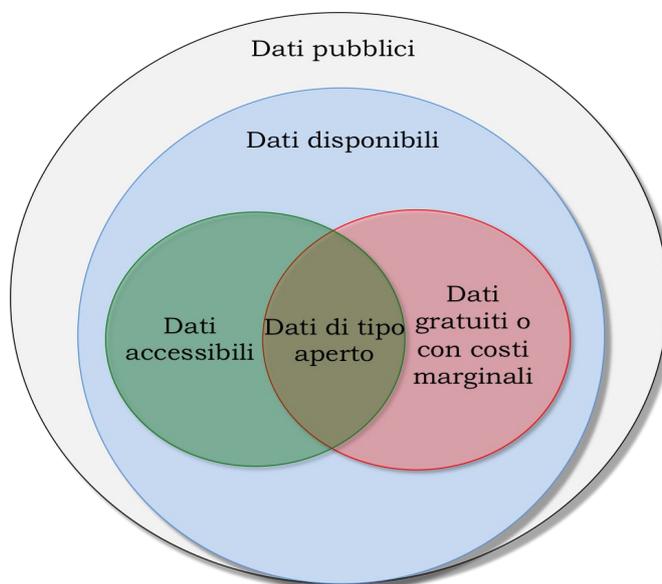
Figura 1 - Tipi di dato pubblico

Per la scelta del formato da utilizzare per il rispetto delle caratteristiche di cui sopra, si può fare riferimento all'allegato A, che riporta il modello a 5 stelle per i dati aperti, e all'allegato B, che descrive i formati più comuni.

REQUISITO 2: dlgs36-2006/opendata/req/common/3stars

I dati DEVONO essere resi disponibili in formato aperto e leggibile meccanicamente ad un livello di **almeno** 3 stelle nella classificazione del modello di cui all'allegato A.

Ulteriori elementi per la scelta del formato da utilizzare si possono rinvenire nella Tabella seguente tratta dal documento "**Data.europa.eu - Data Quality Guidelines**" (v. box "Risorse

utili" alla fine del paragrafo) che riporta l'elenco dei formati utilizzati comunemente specificando quali siano leggibili meccanicamente e non proprietari. L'ultima colonna indica il numero di stelle ottenibili utilizzando il corrispondente formato per la pubblicazione dei dati. I formati evidenziati in verde sono quelli che dovrebbero essere utilizzati; se questo non è possibile allora si possono utilizzare quelli evidenziati in giallo, mentre sono da escludere quelli evidenziati in rosso.

Formato	Non-proprietario	Leggibile meccanicamente	Stelle raggiungibili
RDF	Sì	Sì	★★★★
XML	Sì	Sì	★★★
JSON	Sì	Sì	★★★
CSV	Sì	Sì	★★★
ODS	Sì	Prevalentemente	★★★
XLSX	Sì	Prevalentemente	★★★
XLS	No	Prevalentemente	★★
TXT	Sì	Prevalentemente	★*
HTML	Sì	Prevalentemente	★*
PDF	Sì	No	★
DOCX	Sì	No	★
ODT	Sì	No	★
PNG	Sì	No	★
GIF	No	No	★
JPG/JPEG	No	No	★
TIFF	No	No	★
DOC	No	No	★

Tabella 1 - Formati più comuni per i dati aperti e relativi livelli di apertura

NOTA - ***** Il documento evidenzia che i formati *txt* e *html* dovrebbero essere valutati con tre stelle, poiché i dati potrebbero essere progettati per essere leggibili dalla macchina. Tuttavia, viene assegnata solo una stella perché questi formati non erano originariamente concepiti per rappresentare contenuti leggibili dalla macchina ma solo dall'uomo. La rappresentazione di contenuti leggibili automaticamente in questi formati non soddisfa, quindi, le migliori pratiche e pertanto non sono consigliati dagli autori del documento.

Nella tabella i formati per le immagini (PNG, GIF, JPG/JPEG, TIFF) sono presenti nella zona evidenziata in rosso e, quindi, ne è sconsigliato il loro uso. Tale raccomandazione, però, è da intendere nell'utilizzo di tali formati per rappresentare dati che, invece, andrebbero rappresentati attraverso formati più appropriati (per es. una tabella rappresentata come immagine). Questo implica che per le immagini vere e proprie si possono utilizzare i relativi formati come indicato nell'Allegato B.

REQUISITO 3: dlgs36-2006/opendata/req/common/more-formats

Nel caso in cui un dato sia disponibile in più formati, almeno uno di essi DEVE essere coerente con il **REQUISITO 2**.

Raccomandazione 1: dlgs36-2006/opendata/rec/common/lod

SI RACCOMANDA un percorso graduale verso la produzione nativa di Linked Open Data – LOD (livello cinque stelle nella classificazione del modello di cui all'allegato A).

Nel caso in cui mettere a disposizione i propri dati secondo le indicazioni di cui sopra, per soddisfare richieste di apertura, comporti attività che vanno al di là della semplice manipolazione, e che, quindi, implicherebbero difficoltà sproporzionate, gli enti pubblici (cioè pubbliche amministrazioni e organismi di diritto pubblico) non hanno l'obbligo di adeguare i documenti o crearne nuovi o fornire estratti di documenti, fermi restando i limiti imposti dalla normativa in tema di protezione dei dati personali. In tal caso, attraverso un apposito provvedimento, l'ente titolare DEVE motivare le difficoltà sproporzionate indicando le attività sui dati che eccedono la semplice manipolazione.

REQUISITO 4: dlgs36-2006/opendata/req/common/no-od

Nel caso in cui, per soddisfare richieste di apertura, rendere disponibili i dati per il riutilizzo comporti attività che vanno al di là della semplice manipolazione che implicherebbero difficoltà sproporzionate, il titolare dei dati non ha l'obbligo di adeguare i documenti o crearne nuovi o fornire estratti di documenti, motivando, attraverso un apposito provvedimento, le difficoltà sproporzionate anche indicando le attività sui dati che eccedono la semplice manipolazione.

Secondo quanto indicato nella Direttiva al Considerando (32), la valutazione della proporzionalità dello sforzo deve prendere in considerazione la dimensione e il bilancio di funzionamento dell'ente. Questo implica che potrebbero essere considerate giustificazioni ammissibili la mancanza di personale tecnico (per es. in un piccolo Comune, a meno di un'azione di sussidiarietà da parte di un ente sovraordinato) o di fondi da dedicare alle attività necessarie per assecondare le richieste. Tale indicazione si evince anche dalle disposizioni relative ai dati dinamici di cui all'art. 6, commi 5 e 6 del Decreto, in cui lo sforzo sproporzionato è messo in relazione alle capacità economiche e tecniche dell'ente.

Ciò stante, anche il concetto di "semplice manipolazione" può cambiare in funzione dei due parametri innanzi indicati, se si fa riferimento alla disponibilità nell'ente di risorse umane, finanziarie e strumentali. In altre parole, le operazioni che sono riconducibili alla semplice manipolazione per un ente, per via delle capacità tecniche ed economiche disponibili, possono essere considerate uno sforzo sproporzionato in altri enti che di quelle capacità non dispongono. Per tale motivo, è necessario che vengano indicate puntualmente, con tutte le evidenze del caso, le attività necessarie per rispondere alla richiesta di apertura e che potrebbero rappresentare difficoltà sproporzionate per l'ente.

È bene, infine, rimarcare che il **REQUISITO 4** è applicabile sia ai dati già esistenti che a quelli che saranno prodotti dopo l'entrata in vigore delle presenti Linee Guida, ferma restando la disposizione di cui all'art. 6, comma 2 del Decreto che non prevede un obbligo di creazione di nuovi dati, oltre all'adeguamento di dati già esistenti, nel caso di difficoltà sproporzionate.

Quanto indicato in questo paragrafo è valido per tutte le tipologie di dati.

Il Decreto individua particolari tipi di dati quali i dati dinamici, le serie di dati di elevato valore e i dati della ricerca per i quali, nei paragrafi successivi, saranno fornite specifiche indicazioni supplementari in attuazione di quanto disposto dal Decreto.

Per l'implementazione delle regole tecniche contenute nelle presenti Linee Guida, AgID può adottare, ai sensi dell'art. 3, comma 2 del *Regolamento per l'adozione di linee guida per l'attuazione del Codice dell'Amministrazione Digitale* di cui alla Determinazione AgID n. 160/2018[15], opportune **guide operative** per fornire ai soggetti interessati indicazioni e istruzioni più dettagliate, anche in relazione a specifici casi d'uso.

Risorse utili – par. 4.1

- Open Data Goldbook for Data Managers and Data Holders - *Practical guidebook for organizations wanting to publish Open Data*, European Data Portal, 2018
- *data.europa.eu – Data quality guidelines*, Publications Office of the European Commission, 2021

[15] v. https://trasparenza.agid.gov.it/archivio19_regolamenti_0_5376.html

4.2 Dati dinamici

Il Decreto definisce, all'art. 2, comma 1, lettera c-sexies), **i "dati dinamici"** come "*documenti informatici, soggetti ad aggiornamenti frequenti o in tempo reale, in particolare a causa della loro volatilità o rapida obsolescenza*". La Direttiva ribadisce, inoltre, che "*i dati generati da sensori sono solitamente considerati dati dinamici*" (art. 2, punto 8). Esempi di dati dinamici sono, quindi, i dati ambientali, relativi al traffico, satellitari o meteorologici.

I dati dinamici costituiscono una delle tipologie più preziose dal punto di vista commerciale, in quanto possono essere utilizzati per prodotti e servizi che forniscono informazioni in tempo reale, come app di viaggio o di trasporto. Considerato che **il valore economico di tale tipologia di dati dipende dall'immediata disponibilità dell'informazione e dagli aggiornamenti regolari**, come indicato nel Decreto, i dati dinamici, oltre ad essere coerenti con le indicazioni generali di cui alle presenti Linee Guida, DEVONO essere resi **disponibili per il riutilizzo immediatamente dopo la raccolta tramite API adeguate e, ove possibile, attraverso download in blocco**. Per ulteriori dettagli e indicazioni su tali modalità di pubblicazione dei dati si rimanda al par. **7.1.1**.

REQUISITO 5: dlgs36-2006/opendata/req/dynamic/api
I dati dinamici DEVONO essere messi a disposizione per il riutilizzo attraverso API coerenti con il **REQUISITO 27**.

Il **REQUISITO 27** citato indica, in sintesi, che le API implementate per rendere disponibili i dati dinamici DEVONO essere conformi ai requisiti definiti nell'ambito del modello di interoperabilità ed essere registrate nel catalogo API, reso disponibile attraverso la Piattaforma Digitale Nazionale Dati (PDND) in attuazione dell'art. 50-ter del **CAD**.

Raccomandazione 2: dlgs36-2006/opendata/rec/dynamic/bulk
Ove possibile, opportuno o necessario, SI RACCOMANDA di rendere disponibili i dati dinamici anche attraverso download in blocco.

Nel caso in cui mettere a disposizione tali dati subito dopo la raccolta costituisca uno sforzo sproporzionato per il titolare del dato, a causa di non sufficienti capacità finanziarie o tecniche, il Decreto consente a detto titolare di **posticipare la disponibilità dei dati ad un termine successivo e di applicare restrizioni tecniche temporanee**; detto termine e tali restrizioni tecniche DEVONO essere definiti attraverso un apposito provvedimento in modo, però, da non compromettere lo sfruttamento del potenziale economico e sociale dei dati.

Sul concetto di sforzo sproporzionato, si richiamano anche qui le considerazioni espresse nel paragrafo precedente.

REQUISITO 6: dlgs36-2006/opendata/req/dynamic/release

I dati dinamici DEVONO essere resi disponibili per il riutilizzo immediatamente dopo la raccolta.

Se, per motivi di capacità finanziarie o tecniche, ciò non sia possibile, allora i dati dinamici POSSONO essere resi disponibili per il riutilizzo entro un termine e/o con temporanee restrizioni tecniche, tali, però, da non pregiudicare indebitamente lo sfruttamento del loro potenziale economico e sociale.

Il termine e le restrizioni tecniche di cui sopra DEVONO essere definiti e motivati con apposito provvedimento del titolare del trattamento dei dati.

Il Piano Triennale per l'Informatica nella Pubblica Amministrazione include uno specifico risultato atteso sull'aumento del numero di dati dinamici nell'ambito dell'obiettivo teso a favorire la condivisione dei dati tra le PA e il loro riutilizzo da parte di cittadini e imprese.

Anche ai fini del calcolo dell'indicatore per la misurazione del risultato atteso di cui sopra, nei metadati documentati nel catalogo nazionale di cui all'art. 9 del Decreto è necessario indicare che si tratta di dati dinamici. A tale proposito, indicazioni specifiche saranno incluse nella guida operativa sui cataloghi.

Risorse utili – par. 4.2

- Webinar "Real-time Data", data.europa.eu Academy, 2022
- OGC SensorThings API as an INSPIRE download service

4.3 Serie di dati di elevato valore

La Direttiva e, quindi, il Decreto hanno introdotto la tipologia di **serie di dati denominati "di elevato valore"** definiti come quei "*documenti il cui riutilizzo è associato a importanti benefici per la società, l'ambiente e l'economia, in considerazione della loro idoneità per la creazione di servizi, applicazioni a valore aggiunto e nuovi posti di lavoro, nonché del numero dei potenziali beneficiari dei servizi e delle applicazioni a valore aggiunto basati su tali serie di dati*".

I criteri per l'identificazione di tale tipologia di dati sono indicati all'art. 14, comma 2 della Direttiva, secondo cui deve essere valutata la loro potenzialità:

a) nell'apportare importanti benefici socio-economici o ambientali e servizi innovativi;

b) nel beneficiare un numero elevato di utilizzatori, in particolare PMI;

c) nel contribuire a generare proventi;

d) nell'essere combinati con altre serie di dati.

Al fine di garantire condizioni uniformi di esecuzione della Direttiva in tutti gli Stati membri, l'individuazione delle serie di dati di elevato valore, sulla base dei criteri innanzi citati, è stata effettuata attraverso uno specifico atto di esecuzione di competenza della Commissione Europea, il Regolamento (UE) 2023/138 [**REG-HVD**] (di seguito Regolamento).

Le serie di dati di elevato valore individuate dalla Commissione fanno riferimento alle **categorie tematiche indicate nell'Allegato I della Direttiva**, ovvero:

1. Dati geospaziali
2. Dati relativi all'osservazione della terra e all'ambiente
3. Dati meteorologici
4. Dati statistici
5. Dati relativi alle imprese e alla proprietà delle imprese[16]
6. Dati relativi alla mobilità.

La Direttiva assegna alla Commissione anche la facoltà di individuare categorie aggiuntive a quelle innanzi elencate.

Di seguito vengono definiti requisiti e raccomandazioni sulla base delle prescrizioni più generali presenti nel Decreto e nel Regolamento citato. Tali requisiti e raccomandazioni dovranno essere applicati nei termini temporali indicati nel Regolamento insieme alle specifiche indicazioni riportate in quell'atto. Come previsto dal Regolamento, le disposizioni ivi previste (e di conseguenza requisiti e raccomandazioni definiti nelle presenti Linee Guida) sono applicabili anche alle serie di dati esistenti, create prima della data di applicazione del Regolamento stesso.

In ogni caso, ove si ritenga opportuno o necessario, anche per definire ulteriori indicazioni più dettagliate per la loro implementazione, sarà adottata un'apposita **guida operativa**, quale documento di orientamento a cui si fa cenno nell'art. 5, paragrafo (3), lettera (d) del Regolamento.

[16] Sul tema della proprietà delle imprese, si evidenzia la recente pronuncia della CGE sui "titolari effettivi", che ha dichiarato invalido l'articolo 1, punto 15, lettera c), della direttiva (UE) 2018/843 (antiriciclaggio) e la prevista accessibilità al pubblico delle informazioni sulla titolarità effettiva delle società e delle altre entità giuridiche (v. https://curia.europa.eu/juris/document/document.jsf?text=&docid=268062&pageIndex=0&doclang=IT&mode=req&dir=&occ=first&part=1&cid=7030937). In ogni caso, il Regolamento di esecuzione (UE) 2023/138 chiarisce che le Direttive applicabili sono 2013/34/UE e 2004/109/CE.

Tale guida sarà pubblicata nel sito istituzionale di AgID nonché nei portali dei cataloghi di dati di cui all'art. 9, comma 2 del Decreto.

Al fine di garantire un impatto massimo e agevolarne il riutilizzo, **le serie di dati di elevato valore DEVONO essere rese disponibili gratuitamente** (a meno di alcune eccezioni indicate di seguito nel Requisito 8) e **attraverso API**, oltre che **attraverso download in blocco**.

REQUISITO 7: dlgs36-2006/opendata/req/hvd/api

Le serie di dati di elevato valore DEVONO essere messe a disposizione per il riutilizzo attraverso API coerenti con il **REQUISITO 27** e, se del caso, attraverso download in blocco.

Come già indicato per i dati dinamici, il **REQUISITO 27** citato indica, in sintesi, che le API implementate DEVONO essere conformi ai requisiti definiti nell'ambito del modello di interoperabilità ed essere registrate nel catalogo API, reso disponibile attraverso la Piattaforma Digitale Nazionale Dati (PDND) in attuazione dell'art. 50-ter del **CAD**.

REQUISITO 8: dlgs36-2006/opendata/req/hvd/free

Le serie di dati di elevato valore, individuate dal Regolamento (UE) di esecuzione n. 2023/138, DEVONO essere messe a disposizione gratuitamente ad eccezione di quelle:

- in possesso delle imprese pubbliche qualora ciò determini una distorsione della concorrenza nei pertinenti mercati, se così previsto dagli atti di esecuzione di cui sopra;
- in possesso di biblioteche, ivi comprese le biblioteche universitarie, musei e archivi;
- che, in possesso degli enti pubblici, generano utili per coprire una parte sostanziale dei costi inerenti allo svolgimento dei compiti di servizio pubblico e la loro messa a disposizione gratuita avrebbe un impatto sostanziale sul bilancio di tali enti. In tal caso la possibile esenzione dall'obbligo di mettere a disposizione gratuitamente le serie di dati di valore elevato DEVE durare per un periodo non superiore ai due anni dall'entrata in vigore del Regolamento (UE) di esecuzione n. 2023/138.

L'art. 3 del Regolamento definisce le disposizioni per la pubblicazione applicabili a tutte le categorie di set di dati di alto valore. Il rispetto di **REQUISITO 2**, **REQUISITO 7** e **REQUISITO 27** definiti nelle presenti Linee Guida assicura l'adempimento dei paragrafi (1), (2) e (3) del citato articolo.

Lo stesso articolo dispone inoltre che le amministrazioni titolari di dati di elevato valore debbano designare un punto di contatto per le domande e i problemi relativi alle API al fine di

garantire la disponibilità e il mantenimento delle API stesse e, in definitiva, la pubblicazione regolare ed efficace dei set di dati di alto valore (cfr. art. 3(4)).

Nei metadati che documentano i dati di cui al presente paragrafo, dovrà essere, inoltre, indicato che si tratta di dati di elevato valore (cfr. art. 3(5)). Indicazioni specifiche su come implementare tale disposizione saranno fornite nelle guide operative relative ai profili nazionali di metadati (v. par. **4.6**). Ulteriori indicazioni sui metadati saranno fornite nelle guide operative di cui sopra per adempiere a quanto previsto per i differenti set di dati di elevato valore di cui all'Allegato al Regolamento.

Con riferimento al **REQUISITO 8**, il Regolamento precisa che i dati detenuti dalle imprese pubbliche non rientrano nell'ambito di applicazione del Regolamento stesso, limitandolo, quindi, alle sole Pubbliche Amministrazioni (cfr. Considerando 7).

Il concetto di possesso, indicato al terzo punto del **REQUISITO 8**, è da intendere nell'accezione di titolarità dei dati. A tale proposito si può fare riferimento alla definizione di titolare del dato di cui all'art. 2, comma 1, lettera d) del Decreto, come "*la pubblica amministrazione o l'organismo di diritto pubblico che ha originariamente formato per uso proprio o commissionato ad altro soggetto il documento che rappresenta il dato, o che ne ha la disponibilità*". Altra definizione è fornita all'art. 1, comma 1, lett. cc) del CAD quale "*uno dei soggetti di cui all'articolo 2, comma 2, che ha originariamente formato per uso proprio o commissionato ad altro soggetto il documento che rappresenta il dato, o che ne ha la disponibilità*". Una migliore interpretazione è data nelle Linee Guida RNDT [**LG-RNDT**], secondo cui il titolare del dato è "*la pubblica amministrazione (…) che produce e detiene il dato originale, ovvero la versione di riferimento da cui derivano eventuali copie e che ne può disporre liberamente*". Dal confronto delle due definizioni, si evince anche che quando si parla di "avere la disponibilità" si intende "il disporre liberamente". Estendendo la definizione di cui sopra anche agli altri soggetti destinatari di cui al par. **1.3** (oltre alle pubbliche amministrazioni), è da chiarire che quando, in questo documento, si parla di titolare del/i dato/i ci si riferisce sempre al livello più generale dell'organizzazione, ovvero alla persona giuridica o ente unitario. L'individuazione dei soggetti interni all'organizzazione (persone fisiche, uffici o unità organizzative) responsabili per il processo di apertura dei dati e per l'applicazione delle presenti Linee Guida viene demandata all'autonomia organizzativa dell'ente stesso (v. anche par. **5.1.1**).

Sempre con riferimento al **REQUISITO 8**, se imprese pubbliche, biblioteche, musei ed archivi non sono titolari dei dati ma li gestiscono per conto di una Pubblica Amministrazione, allora il suddetto requisito non è applicabile a questi casi, dovendo essere il possesso riconducibile al titolare.

L’art. 4 del Regolamento definisce le disposizioni per il riutilizzo comuni a tutte le categorie di dati di elevato valore. Il rispetto del **REQUISITO 20** e del **REQUISITO 21** assicura l’adempimento del paragrafo (3) del citato articolo.

In aggiunta alle indicazioni comuni di cui sopra, nell’Allegato I del Regolamento sono individuate, per ciascuna categoria tematica, le serie di dati di elevato valore e sono definite le modalità di pubblicazione e riutilizzo, ivi compresi le condizioni applicabili al riutilizzo, i formati dei dati e dei metadati e le modalità tecniche di diffusione.

Nella tabella che segue sono riportate le disposizioni di pubblicazione e riutilizzo per ciascuna categoria tematica indicate nell’Allegato I al Regolamento e le corrispondenti indicazioni definite nelle presenti Linee Guida, rispettando le quali si assicurano gli adempimenti previsti nel Regolamento stesso. Le indicazioni che non trovano corrispondenza nel presente documento sono indicate in parentesi quadre […] e in corsivo. Tali indicazioni sono comunque vincolanti essendo il Regolamento direttamente applicabile negli Stati membri. Per quanto non riportato nel presente documento si rimanda al Regolamento.

Categoria tematica	Disposizioni Regolamento	Rif. indicazioni LG
Dati geospaziali	Le serie di dati sono messe a disposizione per il riutilizzo: • alle condizioni della licenza Creative Commons BY 4.0 o di una licenza aperta equivalente o meno restrittiva; • in un formato aperto leggibile meccanicamente, pubblicamente documentato e riconosciuto dall’Unione o a livello internazionale; • tramite interfacce di programmazione delle applicazioni ("API") e download in blocco; • [*nella versione più aggiornata*].	REQUISITO 20 REQUISITO 21 REQUISITO 2 REQUISITO 3 REQUISITO 7 REQUISITO 27
	I metadati che descrivono le serie di dati rientranti nell’ambito di applicazione delle categorie tematiche di dati INSPIRE contengono almeno gli elementi di metadati di cui al regolamento (CE) n. 1205/2008 della Commissione.	REQUISITO 16
	Per l'attuazione delle serie di dati sulle parcelle di riferimento e sulle parcelle agricole, gli Stati membri tengono conto dell'attuazione in corso della direttiva 2007/2/CE nonché degli obblighi di cui all'articolo 67, paragrafi 3 e 5, del regolamento (UE) 2021/2116.	REQUISITO 14
Dati relativi all’osservazione della terra e all’ambiente	Le serie di dati sono messe a disposizione per il riutilizzo: - alle condizioni della licenza Creative Commons BY 4.0 o di una licenza aperta equivalente o meno restrittiva; - in un formato aperto leggibile meccanicamente, pubblicamente documentato	REQUISITO 20 REQUISITO 21 REQUISITO 2 REQUISITO 3 REQUISITO 7

	e riconosciuto dall'Unione o a livello internazionale; - tramite API e download in blocco (per le versioni storiche delle serie di dati: API o download in blocco, per quanto possibile e opportuno).	REQUISITO 27
	I metadati che descrivono i dati rientranti nell'ambito di applicazione delle categorie tematiche di dati INSPIRE contengono almeno gli elementi di metadati di cui all'allegato del regolamento (CE) n. 1205/2008.	REQUISITO 16
	Le serie di dati sono descritte in una documentazione online completa e accessibile al pubblico che descrive almeno la struttura e la semantica dei dati.	Par. 5.1.5
	Le serie di dati utilizzano vocabolari e tassonomie controllati, pubblicamente documentati e riconosciuti dall'Unione o a livello internazionale, se disponibili.	Par. 5.1.4
Dati meteorologici	Le serie di dati sono messe a disposizione per il riutilizzo: • alle condizioni della licenza Creative Commons BY 4.0 o di una licenza aperta equivalente o meno restrittiva; • in uno qualsiasi dei formati specificati nella tabella presente al par. 3.2 dell'Allegato I o in un altro formato aperto, leggibile meccanicamente, riconosciuto dall'Unione o a livello internazionale; • tramite API e download in blocco, ad eccezione della serie di dati «dati dei modelli NWP», che è resa disponibile solo tramite API; • [*in base alla frequenza e alla tempestività di aggiornamento specificate nella tabella presente nel par. 3.2 dell'Allegato I*].	REQUISITO 20 REQUISITO 21 REQUISITO 2 REQUISITO 3 REQUISITO 7 REQUISITO 27 Allegato B – 2.4
	I metadati che descrivono la serie di dati sono completi e disponibili sul Web in un formato aperto e ampiamente utilizzato, leggibile meccanicamente.	REQUISITO 15
	Le serie di dati sono descritte in una documentazione online completa e accessibile al pubblico che descrive almeno la la struttura e la semantica dei dati[17].	Par. 5.1.5
Dati statistici	Le serie di dati sono messe a disposizione per il riutilizzo: • [*con la frequenza richiesta dalla legislazione corrispondente di cui al punto 4.1. dell'Allegato I (ad es. mensile, trimestrale, annuale)*]; • alle condizioni della licenza Creative Commons BY 4.0 o di una licenza aperta equivalente o meno restrittiva; • in un formato CSV, XML (SDMX), JSON o in altro formato aperto leggibile	REQUISITO 20 REQUISITO 21 REQUISITO 2 REQUISITO 3 REQUISITO 7 REQUISITO 27 Allegato B – 2.2/2.4

[17] Per esempio, le specifiche INSPIRE su Condizioni atmosferiche ed elementi geografici meteorologici (v. https://inspire.ec.europa.eu/documents/Data_Specifications/INSPIRE_DataSpecification_AC-MF_v2.0.pdf)

	meccanicamente, pubblicamente documentato e riconosciuto dall'Unione o a livello internazionale; • tramite API e download in blocco;	
	I metadati che descrivono le serie di dati sono disponibili sotto forma di file strutturato ben sviluppato contenente almeno una descrizione di dati statistici, concetti statistici, metodologie e informazioni sulla qualità dei dati.	REQUISITO 15[18]
	Le serie di dati sono descritte in una documentazione online completa e accessibile al pubblico che descrive almeno la struttura e la semantica dei dati.	Par. 5.1.5
	Le serie di dati utilizzano vocabolari e tassonomie controllati, pubblicamente documentati e riconosciuti dall'Unione o a livello internazionale, se disponibili.	Par. 5.1.4
Dati relativi alle imprese e alla proprietà delle imprese[19]	Le serie di dati sono messe a disposizione per il riutilizzo: • [*senza indebito ritardo dopo l'ultimo aggiornamento*]; • alle condizioni della licenza Creative Commons BY 4.0 o di una licenza aperta equivalente o meno restrittiva, con condizioni supplementari relative al riutilizzo dei dati personali, se del caso; • in un formato aperto leggibile meccanicamente e riconosciuto dall'Unione o a livello internazionale (XHTML per i documenti che rientrano nell'ambito di applicazione del Regolamento Delegato (UE) 2018/815 della Commissione; altri formati, se e ove prescritto dal diritto dell'Unione applicabile), con metadati completi ([*per i documenti che rientrano nell'ambito di applicazione del Regolamento Delegato (UE) 2016/1437 della Commissione, i metadati specificati in tale regolamento, se del caso;*] per gli altri documenti, i metadati prescritti dal diritto dell'UE applicabile, se del caso); la leggibilità meccanica non è imposta per i dati conservati in formati non leggibili meccanicamente (ad es. documenti e	REQUISITO 20 REQUISITO 21 REQUISITO 2 REQUISITO 3 REQUISITO 15 REQUISITO 7 REQUISITO 27 Allegato B – 2.4

[18] Se per i metadati dei dati statistici è utilizzato SDMX, tramite la specifica StatDCAT-AP (v. https://joinup.ec.europa.eu/collection/semantic-interoperability-community-semic/solution/statdcat-application-profile-data-portals-europe/about) tali metadati possono essere resi disponibili nel profilo DCAT-AP e quindi essere documentati nel portale nazionale dei dati aperti.

[19] Con riferimento ai dati riguardanti la proprietà delle imprese, si tenga conto che la Corte di Giustizia UE è intervenuta sul regime di conoscenza dei cc.dd. "titolari effettivi", dichiarando invalido l'articolo 1, punto 15, lettera c), della direttiva (UE) 2018/843 e la prevista accessibilità al pubblico delle informazioni sulla titolarità effettiva delle società e delle altre entità giuridiche (v. Corte giustizia UE, Grande Sezione, 22/11/2022, n. 37, cause riunite C-37/20 e C-601/20). In ogni caso, il Regolamento di esecuzione (UE) 2023/138 [**REG-HVD**] chiarisce che le Direttive applicabili sono 2013/34/UE e 2004/109/CE.

	conti aziendali scannerizzati) o per i campi di dati non strutturati/non leggibili meccanicamente inclusi come parte di documenti leggibili meccanicamente; • tramite API e download in blocco; • [*a livello di singola azienda*].	
	Le serie di dati sono descritte in una documentazione online completa e accessibile al pubblico che descrive almeno la struttura e la semantica dei dati.	Par. 5.1.5
	Le serie di dati utilizzano vocabolari e tassonomie controllati, pubblicamente documentati e riconosciuti dall’Unione o a livello internazionale, se disponibili, come il Core Business Vocabulary[20].	Par. 5.1.4
Dati relativi alla mobilità[21]	Le serie di dati della rete di trasporto sono messe a disposizione per il riutilizzo • [*immediatamente dopo l’ultimo aggiornamento*]; • alle condizioni della licenza Creative Commons BY 4.0 o di una licenza aperta equivalente o meno restrittiva; • in un formato aperto leggibile meccanicamente e riconosciuto dall’Unione o a livello internazionale; • tramite API e download in blocco; • [*nella versione più aggiornata*].	REQUISITO 20 REQUISITO 21 REQUISITO 2 REQUISITO 3 REQUISITO 7 REQUISITO 27
	I metadati che descrivono le serie di dati sulle reti di trasporto contengono almeno gli elementi di metadati definiti nel Regolamento (CE) n. 1205/2008.	REQUISITO 16
	Le serie di dati sono descritte in una documentazione online completa e accessibile al pubblico che descrive almeno la struttura e la semantica dei dati.	Par. 5.1.5
	Le serie di dati utilizzano vocabolari e tassonomie controllati, pubblicamente documentati e riconosciuti dall’Unione o a livello internazionale, se disponibili.	Par. 5.1.4

Tabella 2 - Corrispondenza tra le disposizioni del Regolamento (UE) n. 2023/138 per i dati di elevato valore e le indicazioni delle LG

4.3.1 Il ruolo dell’Istituto Geografico Militare (IGM)

Con riferimento alle serie di dati di elevato valore, il Decreto, all’art. 12-bis, comma 2, assegna competenze specifiche all’Istituto Geografico Militare (IGM) in relazione ai propri compiti istituzionali e alla produzione dei documenti cartografici dello Stato dichiarati ufficiali dal medesimo Istituto.

[20] v. https://joinup.ec.europa.eu/collection/registered-organization-vocabulary/solution/registered-organization-vocabulary/release/100

[21] Siccome l’Italia non ha recepito la Direttiva 2005/44/CE, non sono applicabili le disposizioni relative ai set di dati sulle vie navigabili interne che quindi non sono prese in considerazione nelle presenti Linee Guida.

Al fine di garantire la qualità dei dati di elevato valore appartenenti alla categoria "Dati geospaziali" di cui all'Allegato I della Direttiva, individuati attraverso gli atti di esecuzione della Commissione Europea, ai sensi del Decreto, l'IGM utilizza i suddetti elementi di informazione, resi disponibili dai titolari coerentemente ai Requisiti 2, 3, 7 e 8 (cioè secondo le caratteristiche dei dati aperti e attraverso API e/o servizi di rete INSPIRE), per aggiornare e produrre i propri dati geospaziali di interesse, ricorrendo a procedure e interventi opportuni per integrarli o convalidarli direttamente mediante il proprio personale tecnico. Nel caso di dati geospaziali prodotti con modalità tecniche diverse da quelle eventualmente indicate negli atti di esecuzione della Commissione Europea, l'IGM contribuisce alla validazione del contenuto.

Ai fini della produzione dei documenti cartografici dello Stato ai sensi della legge 2 febbraio 1960, n. 68 e dichiarati ufficiali dall'Istituto, il Decreto stabilisce che l'IGM acquisisce documenti cartografici o dati geospaziali d'interesse nazionale resi disponibili dagli organismi di diritto pubblico elencati nel paragrafo **1.3**, titolari e responsabili della validazione dei dati originali.

Il Decreto dispone, inoltre, che le società private che riusano i dati geospaziali resi disponibili dall'IGM debbano fornire copia dei documenti derivati che su richiesta, previa verifica, possono essere dichiarati conformi ai requisiti tecnici di qualità o alle specifiche adottate dall'Istituto. Le modalità per l'inoltro delle suddette richieste, nonché, in coerenza con i precipui compiti istituzionali, le specifiche di interesse sono pubblicate sul sito istituzionale dell'Istituto.

Il Decreto prescrive, infine, che i rilevamenti eseguiti, per qualsiasi scopo, sul territorio nazionale da organismi di diritto pubblico o privati, debbano essere comunicati all'IGM ai fini di un eventuale aggiornamento dei documenti cartografici dello Stato.

In tema di rilevamenti, sono fatti salvi gli artt. 7 e 10 della legge 2 febbraio 1960, n. 68 relativamente alla comunicazione obbligatoria all'IGM nei casi specifici indicati dalla norma e del divieto di cedere a terzi i rilevamenti nei casi di cui sopra.

Ai fini dell'attuazione dei compiti istituzionali relativamente alla condivisione dell'informazione geografica, l'Istituto Geografico Militare può promuovere convenzioni, accordi o protocolli d'intesa con altri enti. Nell'ambito di tali convenzioni, accordi o protocolli, l'Istituto può fornire attività di consulenza nell'individuazione dei dati geospaziali che possono essere divulgati con le caratteristiche di tipo aperto e nel monitoraggio a livello nazionale dei dati geospaziali di tipo aperto in riferimento alla divulgabilità degli stessi e alle eventuali implicazioni in termini di riservatezza e sicurezza nazionale.

Sono fatte salve altresì le disposizioni in termini di deposito legale di cui alla legge 15 aprile 2004, n. 106 e al D.P.R. 3 maggio 2006, n. 252. A tale proposito, l'IGM pubblica nel proprio sito

istituzionale le specifiche per la consegna digitale all'Archivio della Direzione Conservatorie, indicando i requisiti tecnici delle pubblicazioni cartografiche e dei dati geospaziali d'interesse, e rilascia la relativa attestazione di consegna.

4.4 Dati della ricerca

I **dati della ricerca** sono definiti dal Decreto come "*documenti informatici, diversi dalle pubblicazioni scientifiche, raccolti o prodotti nel corso della ricerca scientifica e utilizzati come elementi di prova nel processo di ricerca, o comunemente accettati nella comunità di ricerca come necessari per convalidare le conclusioni e i risultati della ricerca*".

Esempi di tale tipologia di dati sono forniti nella Direttiva: statistiche, risultati di esperimenti, misurazioni, osservazioni risultanti dall'indagine sul campo, risultati di indagini, immagini e registrazioni di interviste, oltre a metadati, specifiche e altri oggetti digitali. La Direttiva precisa anche che i dati della ricerca, come da definizione, sono diversi dagli articoli scientifici, in cui si riportano e si commentano le conclusioni della ricerca scientifica sottostante.

REQUISITO 9: dlgs36-2006/opendata/req/research/free
I dati della ricerca DEVONO essere resi disponibili gratuitamente per il riutilizzo per fini commerciali o non commerciali, nel rispetto della disciplina sulla protezione dei dati personali.

I dati da considerare, come da definizione, sono quelli raccolti o prodotti nel corso della ricerca scientifica finanziata con fondi pubblici, utilizzati come elementi di prova nel processo di ricerca o necessari per convalidare le conclusioni e i risultati della ricerca, e resi pubblici, anche attraverso l'archiviazione in una banca dati pubblica gestita a livello istituzionale o su base tematica, da ricercatori, organizzazioni che svolgono attività di ricerca e organizzazioni che finanziano la ricerca.

Nel processo di identificazione dei dati della ricerca, da rendere disponibili per il riutilizzo, è necessario tenere conto della normativa in materia di protezione dei dati personali, degli interessi commerciali, dei diritti di proprietà intellettuale[22] e dei diritti di proprietà industriale[23].

A tale proposito, la Direttiva indica che bisogna considerare il principio "*il più aperto possibile, chiuso il tanto necessario*", ripreso anche nella Raccomandazione (UE) 2018/790, fermo restando che i "*dati personali trattati a fini statistici o di ricerca scientifica non possono essere utilizzati per prendere decisioni o*

[22] v. legge 22 aprile 1941, n. 633
[23] v. decreto legislativo 10 febbraio 2005, n. 30

provvedimenti relativamente all'interessato, né per trattamenti di dati per scopi di altra natura", come stabilito all'art. 105, comma 1, del D. Lgs. 196/2003 [**D-LGS-196-2003**][24].

Se i dati della ricerca sono dati dinamici, dati di elevato valore o dati territoriali, allora si applica anche quanto indicato rispettivamente ai paragrafi **4.2**, **4.3** e **4.5**.

Il Decreto dispone che i dati della ricerca DEVONO **rispettare i requisiti di reperibilità, accessibilità, interoperabilità e riutilizzabilità** che rappresentano i 4 principi del framework FAIR (**F**indable - **A**ccessible - **I**nteroperable - **R**eusable)[25].

Di seguito le indicazioni per adempiere tale disposizione[26].

Findable (Reperibile) - Il primo requisito da rispettare per garantire il riutilizzo dei dati è di rendere i dati reperibili da macchine ed essere umani. Per fare questo, dovrebbero essere resi disponibili i metadati attraverso una risorsa consultabile online e dovrebbe essere assegnato un identificatore persistente a dati e metadati.

Diversi sono i profili di metadati; per i dati della ricerca si fa riferimento anche al progetto OpenAIRE[27] e alla relativa infrastruttura.

Ai fini delle presenti Linee Guida, anche i dati della ricerca DEVONO essere documentati nel portale nazionale dati.gov.it o, nel caso siano anche dati territoriali, nel RNDT secondo il **REQUISITO 29** e il **REQUISITO 30**. In tali casi, DEVONO essere applicati gli standard di cui al **REQUISITO 15** e al **REQUISITO 16**.

REQUISITO 10: dlgs36-2006/opendata/req/research/findable

Per rendere i dati reperibili È RICHIESTO:

- documentare i dati con metadati dettagliati;
- assegnare a dati e metadati un identificatore univoco e persistente (per es. DOI);
- rendere disponibili online i metadati attraverso una risorsa consultabile, come per esempio un catalogo o un repository, seguendo il **REQUISITO 29** e il **REQUISITO 30**;
- specificare l'identificatore del dato nei metadati.

[24] Si veda anche il Considerando 162 del [**GDPR**] sul trattamento di dati personali per finalità statistiche
[25] https://force11.org/info/the-fair-data-principles/
[26] Nella definizione dei requisiti, si è fatto riferimento alla guida "How to make your data FAIR" pubblicata nell'ambito dell'iniziativa OpenAIRE (v. https://www.openaire.eu/how-to-make-your-data-fair) e la checklist "How fair are your data?" creata da Sarah Jones e Marjan Grootveld (v. https://zenodo.org/record/5111307#.YhEfAd_SKCQ).
[27] v. https://www.openaire.eu/

Accessible (Accessibile) - Deve essere possibile ad essere umani e macchine accedere ai dati attraverso protocolli standard e aperti.

REQUISITO 11: dlgs36-2006/opendata/req/research/accessible

Per rendere i dati accessibili È RICHIESTO:

- consentire l'accesso a dati e metadati a partire dall'identificatore univoco e persistente assegnato;
- utilizzare protocolli standardizzati e aperti (per es. https);
- rendere sempre disponibili i metadati anche quando i dati non sono accessibili (o perché sono applicati meccanismi di autenticazione e autorizzazione all'accesso o perché non più disponibili).

Interoperable (Interoperabile) - Dati e metadati devono poter essere combinati con altri dati e/o strumenti. Per questo, devono essere conformi a formati e standard riconosciuti.

REQUISITO 12: dlgs36-2006/opendata/req/research/interoperable

Per rendere i dati interoperabili È RICHIESTO:

- fornire i dati in formato aperto secondo i requisiti definiti nelle presenti Linee Guida;
- utilizzare gli standard pertinenti per i metadati;
- utilizzare, ove possibile, vocabolari controllati, parole chiave, thesauri e ontologie;
- inserire riferimenti qualificati ad altri dati o metadati.

Reusable (Riutilizzabile) - I dati devono essere ben documentati in modo che possano essere interpretati correttamente, replicati e/o combinati anche in contesti diversi. Ai dati, inoltre, bisogna assegnare una licenza chiara e accessibile in modo che si possa capire che tipo di riutilizzo è consentito. Resta fermo il dovuto rispetto dei limiti al riutilizzo derivanti dalla normativa europea e nazionale in materia di protezione dei dati personali.

REQUISITO 13: dlgs36-2006/opendata/req/research/reusable

Per rendere i dati riusabili È RICHIESTO:

- rendere disponibili dati accurati e ben descritti con molti attributi pertinenti;
- assegnare ai dati una licenza d'uso chiara e accessibile secondo il **REQUISITO 20** e il **REQUISITO 21**;
- rendere chiaro come, perché, quando e da chi i dati sono stati creati e processati;
- seguire i pertinenti standard di dominio per dati e metadati.

La necessità che i dati della ricerca seguano e siano conformi ai principi FAIR è ribadita anche dal Programma nazionale per la ricerca (PNR) 2021-2027 (v. box "Risorse utili"), che considera tale conformità come uno degli "strumenti" per l'attuazione della **scienza aperta**. Si legge: "*I dati come tutti i risultati della ricerca devono innanzitutto essere conformi ai criteri FAIR e cioè essere reperibili tempestivamente, accessibili su richiesta, interoperabili e comunque riusabili con le opportune regole, strumenti e risorse. I criteri FAIR sono la base per la fruizione e il riutilizzo di dati di diversa provenienza tematica e metodologica. Fare in modo che la produzione di dati FAIR diventi lo standard di riferimento per i risultati della ricerca finanziata con risorse pubbliche comporta un'innovazione sostanziale nella prassi scientifica sia a livello di produzione dei dati sia a livello della fruizione dei medesimi per elaborare nuove conoscenze. Si tratta di evoluzione delle prassi della ricerca (utilizzo di informazioni esistenti, sviluppo della multidisciplinarietà e della interdisciplinarità, collaborazione formale e informale) validazione e valutazione dei risultati. Essa apre anche nuove prospettive di innovazione tecnologica (acquisizione automatica di dati FAIR ove possibile e sviluppo di servizi digitali per la cura, l'archiviazione, l'accesso, l'analisi e la fruizione di risorse di calcolo)*".

In tema di politiche di accesso aperto, il PNR include il "Piano nazionale per la scienza aperta" (v. box "Risorse utili") che ha tra i suoi obiettivi, tra l'altro, l'accesso aperto, appunto, agli strumenti di produzione della ricerca, ai dati generati e alle pubblicazioni scientifiche nonché il sostegno all'accesso ai dati FAIR.

Un ruolo essenziale nell'attuazione della strategia sulla scienza aperta e delle politiche di accesso aperto è svolto dalle infrastrutture di ricerca, oggetto del "Piano nazionale per le Infrastrutture di Ricerca" (v. box "Risorse utili") - parte integrante, anch'esso, del PNR - a cui si rimanda per tutti gli aspetti che possono avere un impatto sull'applicazione del Decreto con riferimento alle banche dati pubbliche, gestite a livello istituzionale o su base tematica, necessarie per l'archiviazione dei dati della ricerca da rendere disponibili per il riutilizzo.

Raccomandazione 3: dlgs36-2006/opendata/req/research/fair

Ove possibile, SI RACCOMANDA di seguire e applicare i principi FAIR per tutte le tipologie di dati, non solo per quelli della ricerca.

Risorse utili – par. 4.4

- FAIR principles
- The FAIR data principles
- How to make your data FAIR – Guides for Researchers
- How FAIR are your data? Checklist
- Programma nazionale per la ricerca (PNR) 2021-2027,
- Piano nazionale per la scienza aperta (PNSA) 2021-2017, adottato con DM n. 268 del 28/02/2022
- Piano Nazionale Infrastrutture di Ricerca (PNIR) 2021 – 2027, adottato con DM n. 1082 del 10/09/2021

4.5 Dati territoriali

I dati territoriali sono definiti dal **CAD** come "*i dati che attengono, direttamente o indirettamente, a una località o a un'area geografica specifica*" (cfr. art. 1, comma 1, lettera i-sexies)), in linea con l'analoga definizione presente all'art. 3, punto 2) della Direttiva INSPIRE [**INSPIRE-DIR**]. Nel presente documento il termine "dati geospaziali" è utilizzato come sinonimo di "dati territoriali".

Il Decreto non include particolari disposizioni su tali dati, se non il rimando alla disciplina specifica definita in applicazione della Direttiva 2007/2/CE conosciuta come Direttiva INSPIRE [**INSPIRE-DIR**] e recepita in Italia con il decreto legislativo 27 gennaio 2010, n. 32 [**D-LGS-32-2010**] o i requisiti specifici per le categorie nell'ambito delle quali sono individuati dati di elevato valore afferenti al framework di INSPIRE ("Dati geospaziali", "Dati relativi all'osservazione della terra e all'ambiente", "Dati meteorologici", "Dati relativi alla mobilità") previsti nell'ambito dei dati di elevato valore (v. par. **4.3**).

REQUISITO 14: dlgs36-2006/opendata/req/spatial/inspire
Nel caso di dati territoriali, DEVONO essere applicate le regole tecniche adottate nell'ambito del framework definito in applicazione della Direttiva 2007/2/CE e relativo decreto di recepimento, D. Lgs. n. 32/2010, oltre che di tutte le altre norme ad essi collegate.

Nel caso di dati territoriali, quindi, si DEVE fare riferimento ai Regolamenti europei e alle norme nazionali (e relativi documenti tecnici) per le diverse componenti (metadati, dati e servizi di

dati, servizi di rete, condivisione e monitoraggio) dell'infrastruttura istituita e implementata nell'ambito di INSPIRE. In particolare:

- per i metadati: le Linee Guida recanti regole tecniche per la definizione e l'aggiornamento del contenuto del Repertorio Nazionale dei Dati Territoriali [**LG-RNDT**] (coerenti con il Regolamento (CE) n. 1205/2008) e la guida operativa per la compilazione dei metadati RNDT (v. box "Risorse utili"), coerente con la guida tecnica INSPIRE sui metadati[28];
- per i dati: Regolamento (UE) n. 1089/2010 e le specifiche dei dati per ciascun tema INSPIRE (v. box "Risorse utili");
- per i servizi di dati e i servizi di rete: Regolamento (UE) n. 1089/2010, Regolamento (CE) n. 976/2009 e le Linee Guida tecniche INSPIRE sui servizi di rete e sui servizi di dati territoriali (v. box "Risorse utili");
- per il monitoraggio: Decisione di esecuzione (UE) 2019/1372.

A questi si aggiungono norme e specifiche tecniche nazionali di dominio, definite eventualmente anche come estensione delle regole INSPIRE; tra le altre (v. box "Risorse utili"):

- specifiche tecniche per i Database GeoTopografici (DBGT) di cui al Decreto 10/11/2011;
- specifiche tecniche per le reti di sottoservizi e il Sistema Informativo Nazionale Federato delle Infrastrutture (SINFI) di cui al Decreto del Ministro dello Sviluppo Economico dell'11 maggio 2016;
- specifiche tecniche per l'illuminazione pubblica.

Sono fatte salve tutte le altre norme applicabili ai dati territoriali, non esplicitamente citate in questo paragrafo (come la legge n. 132/2016 relativamente al Sistema Informativo Nazionale Ambientale – SINA - e alla rete SINANET o le norme che disciplinano i Sistemi Informativi Territoriali Regionali).

Stante quanto rappresentato innanzi, per l'apertura e il riutilizzo dei dati territoriali si applicano le indicazioni e i requisiti generali (cioè validi per tutte le tipologie di dati) definiti nelle presenti Linee Guida, facendo riferimento, per i formati, a quelli specifici per tale tipologia di dati, alcuni dei quali riportati nell'Allegato B.

Nel caso in cui i dati territoriali siano anche dati dinamici, allora DEVONO essere applicati anche i requisiti di cui al paragrafo **4.2**. Se, invece, rientrano tra le serie di dati di elevato valore, DEVONO essere applicati anche i requisiti di cui al paragrafo **4.3** e le indicazioni presenti nel Regolamento (UE) 2023/138 [**REG-HVD**], per le categorie "Dati geospaziali", "Dati relativi

[28] https://inspire.ec.europa.eu/id/document/tg/metadata-iso19139

all'osservazione della terra e all'ambiente", "Dati meteorologici", "Dati relativi alla mobilità". Se, infine, i dati territoriali sono anche dati della ricerca, sono da applicare i requisiti di cui al paragrafo 4.4.

Risorse utili – par. 4.5

- Guida operativa per la compilazione dei metadati RNDT, Agenzia per l'Italia Digitale
- INSPIRE Data Specifications
- INSPIRE Technical Specifications for the implementation of Network Services
- Technical Guidance for INSPIRE Spatial Data Services
- Regole tecniche nazionali per i dati territoriali (DBGT, SINFI, illuminazione pubblica)

4.6 Metadati

La metadatazione ricopre un ruolo essenziale laddove i dati siano esposti a utenti terzi e a software. I metadati, infatti, consentono una maggiore comprensione e rappresentano la chiave attraverso cui abilitare più agevolmente la ricerca, la scoperta, l'accesso e quindi il riutilizzo dei dati stessi. Per i metadati descrittivi generali, ovvero non dipendenti dalle tipologie di dati, DEVE essere applicato il profilo nazionale DCAT-AP_IT, rispettando le obbligatorietà e le raccomandazioni ivi previsti e seguendo gli esempi definiti nella relativa specifica e ontologia. Il profilo è stato definito con le "**Linee guida per i cataloghi dati**" (v. box "Risorse utili") pubblicate prima delle modifiche apportate al **CAD** dal decreto legislativo 13 dicembre 2017, n. 217 e dal "*Regolamento per l'adozione di linee guida per l'attuazione del Codice dell'Amministrazione Digitale*"[29]. A seguito di tali nuove previsioni normative e regolamentari, dette Linee Guida, potranno essere soggette a revisione con la nuova denominazione "**Guida operativa per i cataloghi dati**". A differenza del presente documento che, ai sensi dell'art. 71 del **CAD**, è aggiornato o modificato secondo la procedura prevista dal medesimo articolo, la guida operativa di cui sopra potrà essere aggiornata o modificata ogni qualvolta sarà necessario.

REQUISITO 15: dlgs36-2006/opendata/req/md/dcat-ap-it

I dati, con esclusione di quelli territoriali, resi disponibili per il riutilizzo DEVONO essere documentati attraverso metadati conformi al profilo DCAT-AP_IT definito con la guida operativa per i cataloghi dati.

[29] v. https://trasparenza.agid.gov.it/archivio19_regolamenti_0_5376.html

Per i dati territoriali, il profilo di metadati da considerare è quello definito con le Linee Guida RNDT [**LG-RNDT**]. Tale profilo è coerente con la Direttiva INSPIRE [**INSPIRE-DIR**] e relativa norma di recepimento [**D-LGS-32-2010**], nonché con il Regolamento (CE) n. 1205/2008 relativo ai metadati.

REQUISITO 16: dlgs36-2006/opendata/req/md/rndt

I dati territoriali resi disponibili per il riutilizzo DEVONO essere documentati esclusivamente attraverso metadati conformi alle "*Linee Guida recanti regole tecniche per la definizione e l'aggiornamento del contenuto del Repertorio Nazionale dei Dati Territoriali*" e le relative guide operative.

Risorse utili – par. 4.6

- Guida operativa per i cataloghi dati
- Guida operativa per la compilazione dei metadati RNDT

Capitolo 5

Aspetti organizzativi e qualità dei dati

Il Decreto non contiene specifiche disposizioni sugli aspetti organizzativi e di qualità dei dati. Si ritiene però, utile, riportare nel seguito alcune indicazioni mirate, alcune delle quali, peraltro, già presenti nelle *Linee Guida per la valorizzazione del patrimonio informativo pubblico*[30], sebbene organizzate diversamente e ove necessario integrate e/o riviste, come raccomandazioni su tali aspetti che si ritengono cruciali per un processo di apertura e di riutilizzo efficace e sostenibile.

I requisiti presenti in questo capitolo sono relativi esclusivamente alle richieste di riutilizzo esplicitamente previste dal Decreto (cfr. art. 5), nella nuova formulazione introdotta con il decreto legislativo n. 200/2021 [**D-LGS-200-2021**].

5.1 Aspetti organizzativi

Il processo di apertura di un dato è frutto di una catena di processi e di una serie di attività di analisi ed elaborazione finalizzate al miglioramento della qualità e dell'accesso al dato stesso.

Nella **Figura 2** che segue è rappresentato un possibile percorso di preparazione dei dati per garantirne la produzione e la pubblicazione e/o messa a disposizione di qualità, necessariamente elastico per garantirne, parimenti, l'applicabilità alle diverse realtà amministrative del territorio nazionale.

La rappresentazione di tale percorso è un adattamento dell'analogo processo definito nel documento "*Data quality guidelines*" dell'Ufficio delle Pubblicazioni della Commissione Europea[31]. Uno schema analogo, sebbene riferito ai *linked open data*, è definito anche nelle *Linee Guida per l'interoperabilità semantica attraverso i Linked Open Data*[32]. A tale proposito, **SI RACCOMANDA di seguire le suddette Linee Guida per il processo di produzione di Linked Open Data**.

In considerazione del possibile diverso punto di partenza del processo di apertura dei dati, alcune fasi indicate nel percorso suddetto possono non essere prese in considerazione. Se, per esempio, l'obiettivo è quello di produrre un nuovo dato - quindi non ancora esistente - le attività di ricognizione e di analisi possono essere by-passate e le prime fasi da considerare potranno essere

[30] https://docs.italia.it/italia/daf/lg-patrimonio-pubblico/it/stabile/index.html
[31] v. box "Risorse utili" al par. 4.1
[32] v. box "Risorse utili" al par. 5.1.4

quella della modellazione e della definizione di priorità e del percorso di apertura (inserita nella fase di identificazione), che rimane comunque valida. Il percorso, inoltre, potrebbe essere modulato anche in forma ciclica, organizzando eventualmente le attività previste in più interazioni.

L'implementazione del processo deve avvenire in maniera costante: **le attività non si esauriscono con la mera pubblicazione dei dati, ma devono prevedere momenti continui di aggiornamento, monitoraggio e coinvolgimento degli utenti finali**, fasi non rappresentate nel percorso in oggetto ma ugualmente importanti per graduare il processo di apertura e pubblicazione dei dati sulle effettive esigenze degli utenti.

- **Identificazione**
  - Ricognizione
  - Analisi dei vincoli
  - Definizione priorità e percorso di apertura
- **Analisi**
  - Analisi della qualità
  - Bonifica
  - Analisi di processo
- **Arricchimento**
  - Vocabolari controllati
  - Mashup
  - Linking nei LOD
- **Modellazione e Documentazione**
  - Schemi dei dati
  - Modelli dati
  - Conservazione e storicizzazione
- **Validazione**
  - Valutazione della qualità dei dati
- **Pubblicazione**
  - Metadatazione
  - Politiche di accesso e licenza
  - Scelta modalità (accesso diretto - portale dati - API - Triple store)

Figura 2 - Processo di preparazione dei dati

Un altro modello metodologico per la gestione dei dati aperti è rappresentato da **Open Data Management Cycle (ODMC)**[33], già utilizzato in alcune Pubbliche Amministrazioni, che include un modello dei processi, un modello informativo e un modello organizzativo. Per tutti gli aspetti che sono allineati alle presenti Linee Guida, il modello indicato può rappresentare un utile riferimento per un approccio comune alla gestione del patrimonio informativo. Nel modello indicato, inoltre, sono presenti indicazioni anche per le fasi non considerate nel processo rappresentato nella **Figura 2**.

Ai fini di valutare l'efficacia del processo di apertura, è anche utile misurare l'impatto dell'utilizzo dei dati aperti. Non esiste una definizione di "**impatto dei dati aperti**" nella legislazione o nella documentazione tecnica nazionale. Esso può essere inteso come l'insieme degli effetti e dei benefici ottenuti direttamente o indirettamente dall'utilizzo e/o il riutilizzo dei dati aperti in specifiche politiche, aree o domini. In linea con il rapporto sulla maturità dei dati aperti

[33] v. https://www.odmc.org/

in Europa pubblicato annualmente[34], i domini rispetto ai quali misurare l'impatto possono essere: società, governo, economia e ambiente.

Di seguito alcune indicazioni per ciascun tipo di impatto come presenti nel rapporto di cui sopra:

- **Impatto sociale** - Valuta la misura in cui i dati aperti hanno un impatto sulle sfide sociali, come l'inclusione dei gruppi emarginati nella società, l'aumento di sensibilizzazione riguardo all'edilizia abitativa nelle aree urbane, e le questioni legate alla salute e al benessere;
- **Impatto politico** - Si concentra sui benefici che i dati aperti hanno in tre ambiti: miglioramento dell'efficienza del governo, miglioramento dell'efficacia del governo e aumento della trasparenza e della responsabilità;
- **Impatto economico** - Considera aspetti quali l'impatto macro e micro-economico e i benefici economici per le amministrazioni pubbliche;
- **Impatto ambientale** - Considera aspetti quali la sensibilizzazione sulla qualità dell'acqua e dell'aria, i livelli di rumore nelle città, i sistemi di gestione dei rifiuti, i sistemi di trasporto rispettosi dell'ambiente.

5.1.1 Ruoli e responsabilità

Per attuare il processo indicato innanzi è necessario definire innanzitutto una chiara *data governance* interna con l'individuazione di ruoli e relative responsabilità e integrare le sue fasi sia verticalmente, rispetto ai processi interni già consolidati, che orizzontalmente rispetto alle necessità delle diverse amministrazioni.

L'autonomia organizzativa delle attività di ciascuna Amministrazione non consente di definire un modello unico per la gestione del processo di apertura dei dati e di implementazione delle indicazioni definite nelle presenti Linee Guida, vista anche la clausola di invarianza finanziaria di cui all'art. 13 del Decreto.

Ogni Amministrazione e, all'interno di essa, ogni unità organizzativa POSSONO individuare e definire **un proprio modello procedurale che individui ruoli e responsabilità dei soggetti coinvolti** sulla base delle disponibilità finanziarie, umane e strumentali e nel rispetto

[34] Open Data Maturity Report (v. https://data.europa.eu/en/publications/open-data-maturity/2022)

della normativa nazionale e unionale in materia di protezione dei dati personali, laddove questi siano coinvolti nelle attività di trattamento.

Per le autonomie locali occorre tenere conto, in particolare, di quanto disposto dall'art. 12 del decreto legislativo 18 agosto 2000, n. 267 e dall'art. 1, comma 85, lettera d) della legge 7 aprile 2014, n. 56, relativamente alla gestione dei sistemi informativi e statistici locali e alla funzione fondamentale di Province e Città Metropolitane in materia di raccolta ed elaborazione dati.

Un ruolo essenziale per l'organizzazione di uffici e risorse per il processo di apertura e di riutilizzo dei dati può essere svolto dall'**ufficio del Responsabile per la transizione digitale** (RTD) istituito con l'art. 17 del **CAD**, che ha la responsabilità della transizione alla modalità operativa digitale e che risponde direttamente all'organo di vertice politico o, in assenza di questo, a quello amministrativo dell'ente.

La Circolare n. 3 del 1° ottobre 2018 del Ministro per la pubblica amministrazione (v. box "Risorse utili"), al fine di garantire la piena operatività dell'Ufficio del RTD, raccomanda di prevedere, nell'atto di conferimento dell'incarico o di nomina, oltre che i compiti espressamente previsti dalle norme vigenti in materia, anche il **potere di costituire gruppi tematici per singole attività e/o adempimenti** come può essere il processo di apertura e pubblicazione dei dati.

Nel caso in cui tale potere non sia assegnato al RTD, la costituzione del gruppo "tematico" può essere direttamente decisa dal vertice politico o quello amministrativo.

In ogni caso, **il RTD DEVE essere comunque coinvolto**, stante il suo ruolo di riferimento del vertice politico e/o amministrativo e di figura trasversale a tutta l'organizzazione con potere di agire su tutte le articolazioni amministrative dell'ente, da una parte, e di punto di contatto con l'Agenzia per l'Italia Digitale e la Presidenza del Consiglio dei Ministri, dall'altra, per le questioni connesse alla trasformazione digitale delle pubbliche amministrazioni, come ribadito anche nella citata Circolare n. 3/2018.

Raccomandazione 4: dlgs36-2006/opendata/req/organization/rtd
SI RACCOMANDA di demandare al Responsabile per la transizione digitale (RTD) il compito di costituire un gruppo di lavoro dedicato al processo di apertura dei dati e all'implementazione delle presenti Linee Guida all'interno dell'organizzazione dell'Ente.

Considerata, inoltre, la disciplina relativa ai dati territoriali, definita da uno specifico framework di interoperabilità a livello europeo (v. par. 4.5), è importante garantire un opportuno raccordo

anche con le figure coinvolte nella politica e nella gestione di tali dati, a partire dai rappresentanti dell'amministrazione nella Consulta Nazionale per l'Informazione Territoriale ed Ambientale (CNITA), di cui all'art. 11 del decreto legislativo n. 32/2010 [**D-LGS-32-2010**], se presenti o, in generale, dai referenti delle infrastrutture di dati territoriali.

Nel processo di apertura e riutilizzo dei dati, ove siano coinvolti dati personali, il Responsabile per la protezione dei dati DEVE essere necessariamente coinvolto, ai sensi dell'art. 38, par. 1 del [**GDPR**].

Visto quanto premesso, nelle presenti Linee Guida si considera il **gruppo di lavoro per i dati aperti come possibile struttura per il governo del processo di apertura dei dati**, rimandando all'autonomia organizzativa di ciascun ente la previsione di eventuali altre strutture e figure, oltre all'individuazione dei compiti da assegnare ad essi.

Raccomandazione 5: dlgs36-2006/opendata/req/organization/wg
SI RACCOMANDA di costituire, all'interno dell'organizzazione dell'Ente, un apposito gruppo di lavoro dedicato al processo di apertura dei dati, anche per l'applicazione delle presenti Linee Guida, prevedendo, ove possibile, le strutture e le figure adatte e necessarie a tale scopo, tenendo in considerazione i referenti dei singoli domini (come le infrastrutture di dati territoriali) e prevedendo altresì il necessario coinvolgimento del Responsabile per la protezione dei dati laddove siano coinvolti dati personali.

All'interno del gruppo di lavoro, o comunque nell'ambito dell'organizzazione dell'ente, è bene prevedere un **responsabile per l'apertura dei dati** e il coinvolgimento dei **responsabili e/o dei referenti tematici** che gestiscono e trattano dati nell'ambito delle singole unità organizzative, oltre alle figure che possano fornire il necessario supporto per l'analisi della qualità dei dati, per la protezione dei dati personali eventualmente presenti, per la definizione delle interfacce di accesso ai dati, per la promozione di applicazioni sviluppate a partire dai dati pubblicati, fornendo anche, laddove necessario, esempi di servizi dimostrativi attraverso cui incentivare il riutilizzo. Alcuni membri del team (per es., esperti di tecnologie web, esperti GIS, esperti di tecnologie e strumenti per i Linked Data) possono occuparsi della gestione del processo di apertura del dato dal punto di vista IT.

È importante garantire il **raccordo e la consultazione con le altre figure coinvolte nel processo di digitalizzazione della pubblica amministrazione**, ovvero responsabile per la conservazione documentale, responsabile per la prevenzione della corruzione e la trasparenza,

responsabile della protezione dei dati, responsabile dei sistemi informativi (se non coincide con il RTD), responsabile per la sicurezza.

L'azione di raccordo e consultazione di cui sopra può essere espletata dal RTD a cui, sulla base della Circolare n. 3/2018 citata innanzi, può essere deputata, nell'atto di nomina, l'adozione dei più opportuni strumenti per garantire tale azione.

Risorse utili – par. 5.1.1

Circolare del Ministro per la pubblica amministrazione n. 3 del 1° ottobre 2018 - Responsabile per la transizione digitale - art. 17 decreto legislativo 7 marzo 2005, n. 82 "Codice dell'amministrazione digitale"

5.1.1.1 Coordinamento tra livello nazionale, regionale e locale

Diverse pubbliche amministrazioni centrali, al fine di adempiere a specifici obblighi normativi a loro assegnati o per dar seguito a impegni presi in iniziative internazionali, hanno necessità di raccogliere dati provenienti dal livello di governo regionale e locale. Viceversa, le amministrazioni centrali e regionali possono detenere dati che sono anche di interesse locale.

In queste situazioni, SI RACCOMANDA alle amministrazioni di coordinarsi tra loro prima di intraprendere iniziative singole isolate. In particolare, le amministrazioni centrali possono assumere un ruolo di coordinamento e di promozione dell'apertura dei dati, definendo anche schemi comuni secondo quanto indicato al par. **5.1.5**. Tale ruolo può essere assunto, in alcuni casi, anche dagli altri Enti sovraordinati (Regioni o Province/Città Metropolitane).

SI RACCOMANDA, quindi, di mantenere il colloquio, mediante scambio di dati, tra il livello centrale, regionale e locale attraverso l'uso dei dati aperti stessi, ove presenti, automatizzando quanto più possibile il processo di acquisizione.

Raccomandazione 6: dlgs36-2006/opendata/rec/organization/coordination

SI RACCOMANDA che le pubbliche amministrazioni statali e regionali che detengono archivi contenenti dati di interesse territoriale – fermo restando il rispetto della disciplina in materia di protezione dei dati personali – prevedano una diffusione dei dati pubblici in forma disaggregata per favorire l'ottimale gestione delle funzioni di raccolta ed elaborazione dati delle Province e Città Metropolitane a supporto di tutti gli enti locali del loro territorio.

5.1.2 Individuazione e selezione

Il processo di apertura dei dati non può che partire da una ricognizione dei dati detenuti e trattati dall'ente e dalla successiva identificazione e selezione di quelli che possono essere resi disponibili per il riutilizzo.

L'apertura può riguardare:

- dati nativi, cioè dati generati dalle amministrazioni nell'adempimento delle proprie funzioni istituzionali;
- dati mashup, cioè dati provenienti da diverse fonti e soggetti a operazioni di integrazione.

In entrambi i casi, l'apertura può essere sollecitata da soggetti esterni interessati attraverso specifiche richieste (v. par. **5.2**).

In questa fase si può fare riferimento ad alcune disposizioni che, in taluni casi, prevedono di rendere disponibili i dati a fini di riutilizzo. In particolare, tra gli altri:

- sulla base della disposizione di cui all'art. 7 del [**D-LGS-33-2013**], "*i documenti, le informazioni e i dati oggetto di pubblicazione obbligatoria ai sensi della normativa vigente, resi disponibili anche a seguito dell'accesso civico di cui all'articolo 5*" del medesimo decreto, fermi restando i limiti e le cautele previsti per i dati personali (v. par. **4.1**);
- ai sensi dell'art. 50-quater del [**CAD**], relativo ai dati generati nella fornitura di servizi in concessione, tutti i dati acquisiti e generati nella fornitura del servizio agli utenti e relativi anche all'utilizzo del servizio medesimo da parte degli utenti, esclusivamente per fini statistici e di ricerca e per lo svolgimento dei compiti istituzionali delle pubbliche amministrazioni.

Per i nuovi dati, che rientrino nell'ambito di applicazione del Decreto, il processo di produzione e apertura DEVE essere guidato dal principio dell'**apertura fin dalla progettazione e per impostazione predefinita**, come, tra l'altro, disposto dall'art. 6, comma 4 del Decreto. A tale proposito, nel par. **6.1.2** sono definite una serie di raccomandazioni per applicare il principio dell'"*open data by design*".

Al riguardo, occorre in ogni caso ricordare con particolare riferimento ai dati personali, che il legislatore europeo nella Direttiva ha evidenziato che "*il riutilizzo dei dati personali è ammissibile soltanto se è rispettato il principio della limitazione della finalità di cui all'articolo 5, paragrafo 1, lettera b) e l'articolo 6 del regolamento (UE) 2016/679*" (Considerando n. 52), come anche esplicitato al par. **4.1**. Pertanto,

come evidenziato anche dal Garante per la protezione dei dati personali, la pubblicazione di informazioni personali online per finalità di trasparenza o di pubblicità dell'azione amministrativa non significa che le stesse costituiscano "dati aperti" né che gli stessi siano, di conseguenza, liberamente riutilizzabili da chiunque e per qualsiasi scopo, bensì impone al soggetto chiamato a dare attuazione agli obblighi di pubblicazione sul proprio sito web istituzionale di determinare – qualora intenda rendere i dati riutilizzabili – se, per quali finalità e secondo quali limiti e condizioni eventuali utilizzi ulteriori dei dati personali resi pubblici possano ritenersi leciti alla luce del "principio di finalità" e degli altri principi di matrice europea in materia di protezione dei dati personali di cui all'art. 5 del [**GDPR**] (cfr. Linee guida del Garante in materia di trasparenza [**LG-GPDP**]).

Quanto al principio di finalità, si ricorda che la normativa europea sulla protezione dei dati personali prevede che tali dati debbano essere raccolti per finalità determinate, esplicite e legittime e che possano successivamente essere trattati solo "*in modo che non sia incompatibile con tali finalità*" (art. 5, par. 1, lett. b del [**GDPR**]). Tale requisito è confermato anche dall' art. 6 del [**GDPR**] laddove è prescritto che il trattamento di dati personali per una finalità diversa da quella per la quale i dati personali sono stati raccolti deve essere "*compatibile*" con la finalità per la quale i dati personali sono stati inizialmente raccolti.

Al fine di verificare la compatibilità della finalità del riutilizzo dei dati personali con quella per la quale i dati personali sono stati inizialmente raccolti è possibile far riferimento ai criteri contenuti nel Regolamento europeo (art. 6, par. 4 del [**GDPR**]), che prevede come il titolare del trattamento debba tenere conto, fra l'altro:

"*a) di ogni nesso tra le finalità per cui i dati personali sono stati raccolti e le finalità dell'ulteriore trattamento previsto;*

b) del contesto in cui i dati personali sono stati raccolti, in particolare relativamente alla relazione tra l'interessato e il titolare del trattamento;

c) della natura dei dati personali, specialmente se siano trattate categorie particolari di dati personali ai sensi dell'articolo 9, oppure se siano trattati dati relativi a condanne penali e a reati ai sensi dell'articolo 10;

d) delle possibili conseguenze dell'ulteriore trattamento previsto per gli interessati;

e) dell'esistenza di garanzie adeguate, che possono comprendere la cifratura o la pseudonimizzazione".

Per meglio individuare per quali scopi ulteriori, compatibili con quelli originari, i dati personali pubblicamente disponibili *online* possono essere utilizzati (o debbano altrimenti essere anonimizzati), si vedano gli elementi condivisi in ambito europeo ed elaborati dal Gruppo Art. 29 nel Parere n. 03/2013 sul principio di limitazione della finalità (v. box "Risorse utili").

Ricognizione - Per i dati nativi, vanno, quindi, individuati, nell'ambito delle strutture organizzative dell'amministrazione, quali dati, tra tutti quelli prodotti, si vogliano rendere aperti, in quanto riutilizzabili da cittadini, imprese e stakeholder in genere, anche per abilitare nuove forme di riutilizzo dell'informazione. Per i dati mashup, le amministrazioni possono raccogliere e integrare informazioni da diverse fonti interne ed esterne che concorrano alla formazione del dato. Per tale tipologia di dati, la parte più importante è la definizione delle modalità di accesso a partire dalle politiche dei singoli produttori dei dati e le relative modalità di rilascio e aggiornamento dei dati stessi.

Analisi dei vincoli giuridici - Alla fase di ricognizione fa seguito l'analisi giuridica delle fonti del dato, fondamentale per garantire la sostenibilità nel tempo del processo di produzione e pubblicazione dei dati considerando i possibili vincoli che possono impedirne o limitarne (anche temporalmente) l'apertura, evidenziando limitazioni d'uso, finalità di competenza, determinazione dei diritti e dei termini di licenza, nonché base giuridica, finalità del trattamento e, in ogni caso, protezione dei dati personali eventualmente presenti.

Per supportare tale analisi, si riporta di seguito una breve "check list" utile alla verifica di alcuni aspetti giuridici da sottoporre a valutazione.

Aspetto	Domanda
Protezione dei dati personali	Nel processo di apertura di atti, documenti e/o dati è stata attentamente verificata l'assenza di informazioni riguardanti persone fisiche identificate o identificabili anche indirettamente, ai sensi dell'art. 4, n. 1) del [**GDPR**], valutati anche gli ulteriori aspetti di cui ai punti riportati di seguito? Più nello specifico: I dati sono liberi da ogni informazione personale che possa identificare in modo diretto l'individuo (nome, cognome, indirizzo, codice fiscale, patente, telefono, email, foto, descrizione fisica, ecc.)? I dati sono liberi da ogni informazione indiretta che possa identificare l'individuo (caratteristiche personali che possono identificare facilmente il soggetto)? I dati sono liberi da ogni informazione sensibile riconducibile all'individuo?

I dati sono liberi da ogni informazione relativa al soggetto che, incrociata con dati comunemente reperibili nel web (ad es. google maps, linked data, ecc.), possa facilmente identificare l'individuo?

I dati sono liberi da ogni riferimento a profughi, protetti di giustizia, vittime di violenze o in ogni caso categorie protette?

I servizi di ricerca sui dati sono tali da poter filtrare i dati in modo da ottenere un solo record geolocalizzato, che sia facilmente e direttamente riconducibile ad una persona fisica?

Nel caso in cui è stata verificata la presenza di dati e informazioni personali, riconducibili in maniera diretta o indiretta a persone fisiche identificate o identificabili, si è provveduto a individuare idonee tecniche di anonimizzazione come indicato nel Parere 05/2014 del Gruppo di lavoro Articolo 29 sulle tecniche di anonimizzazione [**PAR-05-2014**]?

In un'eventuale decisione di rendere riutilizzabili dati personali oggetto di un obbligo di pubblicazione online previsto da un idoneo presupposto normativo ai sensi dell'art. 2-ter, commi 1 e 3, del D. Lgs. 196/2003 (ad es. ai sensi del D. Lgs. 33/2013) o di accogliere eventuali richieste di riutilizzo degli stessi da parte di terzi, è stato valutato se, per quali finalità ed entro quali limiti e condizioni siano eventualmente leciti e non incompatibili utilizzi ulteriori dei dati personali resi pubblici (con esclusione dei dati di cui agli artt. 9-10 del [**GDPR**]), sulla base del principio di limitazione della finalità?

È stata effettuata la valutazione d'impatto in materia di protezione dei dati, ai sensi dell'art. 35 del [**GDPR**], al fine di ridurre il rischio di perdere il controllo sulle medesime informazioni o di dover far fronte a richieste di risarcimento del danno da parte degli interessati?

Nella predetta analisi di impatto sono state effettuate tutte le valutazioni (anche inerenti alla tipologia di licenza)

	indicate dal Garante per la protezione dei dati personali nei propri provvedimenti[35]?
	Si è provveduto a coinvolgere il Responsabile per la protezione dei dati personali, la cui nomina è obbligatoria in ogni PA e nei soggetti indicati dall'art. 37 del [**GDPR**]?
Proprietà intellettuale della sorgente	l'ente è proprietario dei dati, anche se non sono stati creati direttamente da suoi dipendenti?
	sei sicuro di non usare dati per i quali vi è una licenza o un brevetto di terzi?
	se i dati non sono del tuo ente, hai un accordo o una licenza che ti autorizzi a pubblicarli?
Licenza di rilascio	stai rilasciando i dati di cui possiedi la proprietà accompagnati da una licenza?
Limiti alla pubblicazione	hai verificato che non vi siano impedimenti di legge o contrattuali per la pubblicazione dei dati?
	hai verificato che i dati e i documenti non siano tra quelli esclusi dall'applicazione del decreto legislativo n. 36/2006?
Segretezza	hai verificato se non vi siano motivi di ordine pubblico o di sicurezza nazionale, motivi legati al segreto d'ufficio o al segreto di stato o al segreto statistico che ti impediscono la pubblicazione dei dati?
Temporalizzazione	i dati sono soggetti per legge a restrizioni temporali di pubblicazione?
Trasparenza	i dati hanno dei divieti di legge o giurisprudenziali che impediscono la loro indicizzazione da parte di motori di ricerca?

[35] Si veda in particolare: "*Linee guida in materia di trattamento di dati personali, contenuti anche in atti e documenti amministrativi, effettuato per finalità di pubblicità e trasparenza sul web da soggetti pubblici e da altri enti obbligati*" (provv. n. 243 del 15/5/2014, in www.gpdp.it, doc. web n. 3134436) e successivi provvedimenti di aggiornamento.

fermi restando i limiti e le garanzie per il trattamento dei dati personali, i dati rientrano nella lista dell'allegato A del d.lgs. 33/2013 [**D-LGS-33-2013**]?

L'analisi dei vincoli e, quindi, la verifica sulla possibilità di rendere disponibili i dati per il riutilizzo possono essere ulteriormente supportate prendendo in considerazione l'elenco dei documenti esclusi dall'applicazione del Decreto riportato nel par. **1.2**. Se il dato o documento rientra tra uno di quelli dell'elenco di cui sopra, è da escludere dai dati da rendere disponibili come dati aperti. In caso contrario, qualora il dato sia ammesso al riutilizzo, il titolare dei dati deve effettuare una valutazione d'impatto sulla protezione dei dati personali, avendo cura di garantire una protezione dei dati *by design* e *by default*.

Definizione della priorità e percorso di apertura - La ricognizione dei dati consente di avere contezza del patrimonio informativo, mentre l'analisi dei vincoli consente di fare una prima selezione dei dati che possono essere resi disponibili per il riutilizzo. Per i dati non soggetti a vincoli, è necessario individuare criteri e modalità di apertura anche in base ad una eventuale **scala di priorità** basata, per esempio, su un approccio di tipo "*demand-driven*" che tenga conto dell'impatto economico e sociale nonché del livello di interesse e delle necessità degli utilizzatori.

Tra i criteri da considerare per definire la priorità nell'apertura dei dati possono essere considerati:

- la tipologia di dato cui sia attribuito dal Decreto un enorme potenziale economico (dati dinamici, serie di dati di elevato valore e dati della ricerca) per come definita dal Decreto medesimo;
- l'esistenza di richieste pervenute dai riutilizzatori sulla base dell'art. 5 del Decreto e della procedura di cui al par. **5.2**;
- l'esistenza di specifiche disposizioni normative che prevedano di rendere disponibili obbligatoriamente, come dati di tipo aperto, talune tipologie di dati (v. i casi presentati all'inizio di questo paragrafo).

Tale attività potrebbe essere sistematizzata attraverso la definizione di un **percorso di apertura dei dati** da inserire nel Piano Triennale ICT che ciascuna amministrazione, secondo la roadmap definita dalle Linee d'Azione nel Piano triennale nazionale e le modalità operative fornite da AgID, è chiamata a definire anche utilizzando il format PT reso disponibile da AgID stessa (v.

box "Risorse utili"). La predisposizione di tale Piano rientra tra i compiti che la Circolare n. 3/2018 del Ministro per la pubblica amministrazione raccomanda di assegnare al RTD.

Raccomandazione 7: dlgs36-2006/opendata/rec/organization/dmp
SI RACCOMANDA di definire un percorso di apertura dei dati da inserire nel Piano Triennale ICT della singola Amministrazione, la cui definizione può rientrare nei compiti da assegnare al RTD. Tale percorso potrà essere basato su una scala di priorità nell'apertura tenendo in considerazione gli obblighi derivanti dall'applicazione del Decreto per alcune specifiche tipologie di dati, fermo restando il rispetto della normativa unionale e nazionale in materia di protezione dei dati personali. Potranno essere considerate anche specifiche regole di politiche dei dati aperti (Data policy), purché non in contrasto con le strategie e le politiche nazionali.

A tale proposito, si fa presente che le linee d'azione del Piano Triennale nazionale relative al capitolo sui dati includono, tra l'altro:

- l'individuazione dei dataset di tipo dinamico da rendere disponibili in open data in coerenza con quanto previsto dalla Direttiva e la loro documentazione nel catalogo nazionale dei dati aperti;
- la messa a disposizione dei dati territoriali attraverso i servizi di rete di cui a [**INSPIRE-DIR**];
- la documentazione dei dati di tipo aperto attraverso il catalogo nazionale dei dati aperti (in corrispondenza del risultato atteso relativo all'aumento del numero di dataset di tipo aperto).

I criteri di prioritizzazione indicati innanzi, quindi, troverebbero fondamento anche nelle azioni che gli enti devono implementare per raggiungere gli obiettivi e i risultati attesi prefissati dal Piano Triennale nazionale.

Risorse utili – par. 5.1.2

- *Linee guida in materia di trattamento di dati personali contenuti anche in atti e documenti amministrativi effettuato da soggetti pubblici per finalità di pubblicazione e diffusione sul web*, adottate con Deliberazione n. 243 del 15 maggio 2014 del Garante per la Protezione dei Dati Personali
- *Linee Guida recanti indicazioni operative ai fini della definizione delle esclusioni e dei limiti all'accesso civico di cui all'art. 5 co. 2 del D. Lgs. 33/2013* adottate con Determinazione n. 1309 del 28 dicembre 2016 dell'Autorità Nazionale Anticorruzione
- *Guida per la redazione format del Piano Triennale per le pubbliche amministrazioni*, Agenzia per l'Italia Digitale
- *Parere n. 03/2013 sul principio di limitazione della finalità*, Gruppo di Lavoro Articolo 29

5.1.3 Analisi

Analisi della qualità dei dati - All'analisi giuridica delle fonti segue l'analisi della qualità dei dati. Per la definizione del concetto di qualità dei dati si può ricorrere alla norma ISO/IEC 25012, secondo cui "***la qualità dei dati è il grado in cui le caratteristiche dei dati soddisfano esigenze espresse e implicite quando utilizzati in specifiche condizioni***". Nella sezione dedicata alla "qualità dei dati" (v. par. 5.3) si identificano alcune misure e un metodo di valutazione, basati sugli standard ISO di riferimento.

Bonifica - Generalmente, l'analisi della qualità del dato può richiedere una fase di bonifica che si sostanzia in attività di miglioramento di detta qualità che può essere ottenuto eliminando errori e criticità attraverso processi basati sui dati, tramite il confronto con il mondo reale o il confronto incrociato (matching) con altri dataset, oppure adottando azioni di bonifica basate su processi che hanno la caratteristica di analizzare le cause che hanno portato alla scarsa qualità del dato, in modo anche da rivedere gli stessi processi di produzione del dato per garantirne la qualità nel tempo.

Analisi di processo, (re)ingegnerizzazione dei processi organizzativi e produzione dei dati - Ogni dato ha un proprio ciclo di vita, caratterizzato da uno specifico tasso di aggiornamento o manutenzione. Lo Standard UNI CEI ISO/IEC 25024:2016 "Misurazione della qualità del dato"

riporta un esempio dell'intero ciclo di vita del dato composto dalle fasi di: progettazione, acquisizione, integrazione con altri dati, elaborazione, memorizzazione, uso, cancellazione.

Risulta quindi necessario **analizzare il processo organizzativo che produce e gestisce il dato per fare in modo che la produzione di quel dato sia consolidata e diventi stabile, secondo la frequenza di aggiornamento e le modalità di rilascio adottate**.

È preferibile che l'aggiornamento sia operato dal titolare del dato nativo, di prima produzione, possibilmente in maniera coordinata con le altre strutture organizzative dell'ente ovvero degli altri enti competenti per materia e, in generale, altre organizzazioni, anche al fine di evitare duplicazioni.

A tale proposito, a livello territoriale, si deve tenere conto delle funzioni di raccolta ed elaborazione dati che la legge 7 aprile 2014, n. 56, all'art. 1, comma 85, lettera d) riconosce come funzione fondamentale delle Province e delle Città Metropolitane.

5.1.4 Arricchimento

Una volta che i dati siano stati bonificati, possono essere arricchiti attraverso l'integrazione con altri dati e il linking esterno. L'arricchimento è definito dal documento "Data quality guidelines" del Publications Office[36] come risultante del collegamento, ai dati esistenti, dei dati da fonti esterne, ed è un processo che, secondo il documento medesimo, include sia la standardizzazione che specifiche attività di arricchimento vero e proprio.

Chiaramente parlando di arricchimento ci si riferisce a dati già esistenti; le azioni indicate, però, potrebbero essere applicate ai dati già al momento della loro formazione, costituendo buone pratiche da seguire in quella fase.

Vocabolari controllati - Come indicato nel documento "Data quality guidelines", un livello più alto di standardizzazione può essere raggiunto facendo riferimento a vocabolari controllati RDF, quali elenchi di codici, tassonomie, classificazioni o terminologie, definiti nel Piano Triennale ICT 2017-2019 come "*un modo comune e condiviso per organizzare codici e nomenclature ricorrenti in maniera standardizzata e normalizzata*". I vocabolari controllati assegnano a ogni concetto un identificatore univoco e persistente (URI) in modo che quel concetto venga referenziato in maniera non ambigua

[36] v. box "Risorse utili" al par. 4.1

e garantiscono, inoltre, la gestione in modo coerente delle diverse versioni. Possono essere associate, oltre alle etichette, definizioni e descrizioni anche in diverse lingue.

I vocabolari fanno sì che invece di utilizzare nei dati le etichette, queste possano essere referenziate dagli identificatori univoci assegnati, in modo che, se le etichette dovessero cambiare, il riferimento non deve essere adeguato, riducendo l'onere di manutenzione per titolari e fruitori di dati. Considerato, inoltre, che gli URI possono essere dereferenziati (v. par. **7.1.3**), l'etichetta può essere risolta in qualsiasi lingua.

In caso di dati di elevato valore (v. par. **4.3**), per specifiche categorie tematiche (ovvero osservazione della terra e ambiente, statistica, imprese e proprietà delle imprese, mobilità) il Regolamento (UE) [**REG-HVD**] indica esplicitamente che i set di dati DEVONO utilizzare, ove possibile, vocabolari controllati e tassonomie documentati pubblicamente e riconosciuti nell'Unione o a livello internazionale.

I vocabolari controllati sono disponibili nel Catalogo Nazionale della semantica dei dati e, per i dati territoriali, nel Sistema di Registri INSPIRE Italia (v. box "Risorse utili").

Integrazione con altri dati - L'arricchimento dei dati può essere ottenuto, come detto, anche integrando informazioni da sorgenti esterne rendendo in questo modo i dati di origine più significativi e fruibili. Il valore aggiunto è ottenuto, per esempio, producendo i dati mashup già citati in precedenza.

Linking nei LOD - Come detto, il collegamento (linking) dei dati può aumentarne il valore creando nuove relazioni e consentendo così nuovi tipi di analisi.

Nel caso in cui il processo sia finalizzato alla produzione di linked open data, come evidenziato nelle già citate Linee Guida per l'interoperabilità semantica attraverso i Linked Open Data, "*il linking è una funzionalità molto importante e di fatto può essere considerata una forma particolare di arricchimento. La particolarità consiste nel fatto che l'arricchimento avviene grazie all'interlinking fra dataset di origine diversa, tipicamente fra amministrazioni o istituzioni diverse, ma anche, al limite, all'interno di una stessa amministrazione*". Tale collegamento è possibile soprattutto attraverso l'uso coerente di identificatori univoci, gli URI, di cui si è parlato prima e che vengono approfonditi ulteriormente nel par. **7.1.3**.

I linked open data, oltre agli URI, si basano su diversi standard, tra cui RDF, e spesso usano vocabolari controllati RDF per rappresentare terminologia controllata del dominio applicativo di riferimento. Nell'allegato B sono riportati, nella prima parte, anche i principali standard di riferimento necessari anche ad abilitare i livelli 4 e 5 del modello dei dati di cui

all'allegato A. Utilizzando il framework RDF, si può costruire un grafo semantico, noto anche come grafo della conoscenza, che può essere percorso dalle macchine risolvendo, cioè dereferenziando, gli URI HTTP.

Dai dati così disponibili è possibile estrarre automaticamente informazione e derivare, quindi, contenuto informativo aggiuntivo (inferenza) attraverso ragionatori automatici e query SPARQL (v. le Linee Guida indicate nel box "Risorse utili").

Alcune delle fasi indicate nel percorso di cui alla **Figura 2** possono essere comuni al processo di produzione dei Linked Open Data. Tale processo è caratterizzato da altre specifiche fasi, non rappresentate nel percorso comune, ma dettagliate nelle Linee Guida citate innanzi a cui si rimanda, come già raccomandato in precedenza.

Risorse utili – par. 5.1.4

- Linee Guida per l'interoperabilità semantica attraverso i Linked Open Data
- Catalogo nazionale della semantica dei dati
- Sistema di Registri INSPIRE Italia

5.1.5 Modellazione e documentazione

Una fase molto importante nel processo di preparazione dei dati è la definizione di sintassi (cioè struttura) e semantica (cioè contenuto). Questo, oltre a migliorare l'interoperabilità, la qualità e a facilitarne l'elaborazione, aumenta anche il valore dei dati stessi, poiché l'interpretazione errata dei dati diventa meno probabile quando viene fornito il contesto.

Come detto innanzi, questa fase può corrispondere alla prima del processo nel caso in cui il dato ancora non esista e, quindi, si parte dalla definizione di un'ontologia o comunque di un modello dati del dominio, cioè dalla definizione sintattica e semantica dei dati in termini di entità rappresentate, loro attributi e associazioni (cosiddetta fase di modellazione).

Rientra in questa fase anche la documentazione delle modifiche e degli aggiornamenti dei dati oltre alla gestione delle relative versioni e la storicizzazione.

Schemi dei dati – Per facilitare la corretta interpretazione dei dati da parte del fruitore ed evitare qualsiasi ambiguità nella loro comprensione, è necessario condividere anche gli schemi dei dati,

eventualmente specificati separatamente, che forniscono la descrizione di sintassi e struttura dei dati stessi.

A titolo di esempio, con riferimento ad alcuni formati riportati al par. 2 dell'Allegato B (CSV, JSON, XML), gli schemi che potrebbero essere utilizzati sono i seguenti:

- JSON Schema[37] per il formato JSON;
- XSD (XML Schema Definition)[38] per il formato XML;
- Table Schema per il formato CSV definito nell'ambito del toolkit Frictionless Data[39].

Il toolkit Frictionless include anche il cosiddetto Data Package, un "formato contenitore" composto dai metadati che descrivono la struttura e il contenuto del pacchetto (memorizzati in un "descrittore") e le risorse, come i file di dati, che costituiscono il contenuto del pacchetto.

In generale, i metadati di contenuto potrebbero seguire le ontologie nazionali disponibili nel Catalogo nazionale della semantica dei dati, soprattutto per tutti quei dati indipendenti dal dominio specifico.

Modelli dati - Il Piano Triennale ICT 2017-2019 definisce un'ontologia o un modello dati condiviso come "*una concettualizzazione esaustiva e rigorosa nell'ambito di un dato dominio*".

Anche per garantire la coerenza tra i documenti, si richiama qui la fase denominata "Semantica" nel processo digitale individuato e descritto nelle Linee Guida per l'interoperabilità tecnica [**LG-INT**], in cui si evidenzia che la comunicazione tra soggetti DEVE utilizzare modelli dati condivisi, in modo da razionalizzare e uniformare la rappresentazione dell'informazione quale presupposto per favorire l'interoperabilità tra soggetti differenti.

Le Linee Guida di cui sopra hanno già definito una serie di requisiti in tema di modelli dati. Come indicato nel documento citato, pertanto, nell'individuazione delle entità da condividere i diversi soggetti DEVONO:

1) individuare i domini di interesse e in essi determinare le entità da rappresentare in termini di proprietà che li caratterizzano;

2) verificare la presenza delle entità per dominio tra quelli definiti a livello nazionale nella rete di ontologie e vocabolari pubblicati nel Catalogo Nazionale per la semantica dei dati.

[37] v. https://json-schema.org/
[38] v. https://www.w3.org/TR/xmlschema11-1/ e https://www.w3.org/TR/xmlschema11-2/
[39] v. https://frictionlessdata.io/

A integrazione di quanto sopra, considerato quanto previsto dall'art. 6, comma 9 del Decreto, i modelli dati da considerare per i dati territoriali sono quelli definiti nell'ambito delle attività di regolamentazione derivanti dalla Direttiva INSPIRE **INSPIRE-DIR** e nell'ambito del framework nazionale che fa riferimento ai decreti 10/11/2011 e alle attività di estensione delle regole INSPIRE (v. par. **4.5**).

Si aggiunge, pertanto, un ulteriore requisito:

3) nel caso di dati territoriali, verificare la presenza delle entità per dominio tra quelli definiti a livello europeo e nazionale nell'ambito della regolamentazione INSPIRE e la sua estensione nazionale.

Come indicato nelle citate Linee Guida, successivamente all'attuazione delle regole 2) e 3) ci si può trovare in uno dei seguenti casi:

a) tutte le entità e le relative proprietà trovano copertura;

b) almeno una delle entità non è compresa nelle rappresentazioni;

c) almeno una proprietà di un'entità presente non risulta rappresentata.

Nel caso a), il soggetto ha tutti gli elementi per rappresentare il proprio modello dati; viceversa, nei casi b) e c), la stessa amministrazione, in accordo con AgID, valuta l'opportunità di estendere il modello dati a livello nazionale.

La regola che, in generale, deve guidare, è di esaminare modelli dati, ontologie e vocabolari controllati esistenti per verificare se i concetti siano già provvisti di entità, proprietà e, ove presenti, URI ampiamente adottati, specie se in ambito europeo. Solo in caso contrario, l'ente che pubblica i dati può definire e pubblicare, secondo le regole indicate innanzi, il proprio modello dati, ontologia o vocabolario controllato al fine di definire concetti che non siano stati specificati altrove.

In caso di dati di elevato valore (v. par. **4.3**), per tutte le categorie tematiche, tranne quelle relative a dati per i quali si applicano i modelli dati INSPIRE, il Regolamento UE indica esplicitamente che i set di dati DEVONO essere descritti in una documentazione online completa e pubblicamente disponibile che contenga almeno la definizione della struttura e della semantica dei dati.

Conservazione e storicizzazione - I dataset rilasciati costituiscono non solo una risorsa per la collettività, ma un prezioso patrimonio anche per le pubbliche amministrazioni che possono in questo modo archiviare in modo alternativo i loro dati in modalità indipendente dagli applicativi

software originali che li hanno prodotti. Per questo motivo è importante premunirsi di un sistema di archiviazione/conservazione che mantenga le diverse versioni dei dati nel lungo periodo. A tal fine si raccomanda di assicurare che le versioni stesse siano accessibili a un URL stabile, che sia anche documentato unitamente alla pubblicazione del dato.

A tale proposito, il Decreto stabilisce che le pubbliche amministrazioni e gli organismi di diritto pubblico debbano utilizzare le modalità per facilitare la conservazione dei documenti disponibili per il riutilizzo secondo quanto previsto dall'articolo 44 del **CAD**.

REQUISITO 17: dlgs36-2006/opendata/req/organization/rndt
Le pubbliche amministrazioni e gli organismi di diritto pubblico, al fine di facilitare la conservazione dei documenti disponibili per il riutilizzo, DEVONO seguire le indicazioni fornite nelle "*Linee Guida sulla formazione, gestione e conservazione dei documenti informatici*".

Stante quanto indicato nel par. **4.1** relativamente alla distinzione tra documenti e dati, il **REQUISITO 17** si applica solo ai documenti che rientrano nell'ambito di applicazione delle Linee Guida richiamate nel requisito stesso.

Risorse utili – par. 5.1.5

- Catalogo nazionale per la semantica dei dati
- *Regolamento sui criteri per la fornitura dei servizi di conservazione dei documenti informatici*, adottato con la Determinazione AgID n. 455/2021 del 25 giugno 2021

5.1.6 Validazione

La validazione dei dati è una parte essenziale di qualsiasi processo di apertura, comunque propedeutica alla pubblicazione e al riutilizzo. Essa viene definita come "*un'attività volta a verificare se il valore di un dato proviene dall'insieme dato (finito o infinito) di valori accettabili*" o "*come un processo che assicura la corrispondenza dei dati finali (pubblicati) con una serie di caratteristiche qualitative*"[40].

[40] Definizioni riportate nel documento "Methodology for data validation 2.0" (v. https://ec.europa.eu/eurostat/ramon/statmanuals/files/methodology_for_data_validation_v2_0_rev2018.pdf)

In sintesi, lo scopo della validazione dei dati è quello di assicurare un certo livello di qualità ai dati stessi.

Di analisi di qualità dei dati si è già accennato nella fase post ricognizione dei dati e, come già innanzi, si rimanda al par. 5.3 dedicato alle caratteristiche e alle misure della qualità.

La differenza e le relazioni tra validazione e qualità sono indicate nel documento "Methodology for data validation 2.0" (v. box "Risorse utili"): la validazione dei dati si focalizza sulle dimensioni della qualità relative alla "struttura e al contenuto dei dati", ma non sugli aspetti della qualità che riguardano i processi. Si può dire, quindi, che la validazione è relativa solo alle caratteristiche di qualità definite "inerenti" nello Standard ISO/IEC 25012:2008 e riportate nel citato par. 5.3.

La fase di validazione può essere un ulteriore passaggio per la verifica dei dati a valle delle altre operazioni effettuate prima della pubblicazione, per es. per l'arricchimento. Può essere anche intesa come la prima e unica fase del processo per la verifica della qualità nel caso di nuovi dati che, quindi, non rientrerebbero nella fase di ricognizione e nella conseguente analisi.

Il documento "Methodology for data validation 2.0" individua sei livelli di validazione che si riportano di seguito rimandando al documento citato per ulteriori approfondimenti.

- **Livello 0**: coerenza con i requisiti strutturali IT previsti;
- **Livello 1**: coerenza all'interno del set di dati;
- **Livello 2**: coerenza con altri set di dati all'interno dello stesso dominio e all'interno della stessa origine dati;
- **Livello 3**: coerenza all'interno dello stesso dominio tra origini dati diverse;
- **Livello 4**: coerenza tra domini separati dello stesso fornitore di dati;
- **Livello 5**: coerenza con i dati di altri fornitori di dati.

Risorse utili – par. 5.1.6

📖 *Methodology for data validation 2.0*, Essnet Validat Foundation, Revised edition 2018

5.1.7 Pubblicazione

Prima di procedere alla pubblicazione è necessario creare i metadati, definire le politiche di accesso e il modello di licenza da applicare e identificare i canali per la pubblicazione, anche in base

alla tipologia dei dati, tra accesso diretto (del singolo dataset e/o in blocco), portale dati, API e, in caso di linked open data, triple store.

Nella scelta del canale si devono tenere in considerazione anche i requisiti definiti nelle presenti Linee Guida; ad esempio, per come disposto dal Decreto, i dati dinamici e le serie di dati di elevato valore DEVONO essere resi disponibili attraverso API e, ove possibile, attraverso download in blocco, il che non significa, naturalmente, che non possano essere resi disponibili anche attraverso altri canali ma che quella individuata in prima battuta possa essere l'opzione preferenziale. Il **Capitolo 7** fornisce indicazioni, raccomandazioni e ulteriori elementi per la pubblicazione dei dati.

Metadatazione - Come detto precedentemente, la metadatazione è cruciale: i metadati certificano le caratteristiche del dato. Si ricorda, a tale riguardo, di seguire i profili di metadati indicati nel paragrafo **4.6** a cui si rimanda e che consentono di specificare i più importanti metadati descrittivi per i dataset (per es., soggetti e relativi ruoli, contestualizzazione geografica e temporale, licenza, frequenza di aggiornamento, aspetti di distribuzione, punto di contatto, ecc.).

Politiche di accesso e licenza – Nell'ambito del processo di apertura sono da tenere in considerazione eventuali forme di aggregazione dei dati e restrizioni di accesso, che hanno anche un impatto sulla scelta della licenza, della quale si tratterà al **Capitolo 6**, e sulla protezione dei dati personali. Sebbene sia sconsigliato restringere l'accesso ai dati o procedere con la pubblicazione di aggregazioni degli stessi (in generale non è opportuno che l'esposizione del dato lavorato avvenga senza che sia stato pubblicato prioritariamente il dato grezzo), esistono casi in cui i dati possono essere diffusi solo in forma anonima (ad esempio i redditi), ossia a un livello di aggregazione tale da impedire di identificare le persone cui i dati si riferiscono (cfr. indicazioni in merito presenti nei paragrafi **4.1** e **5.1.2**).

A tal fine, è bene definire delle politiche di accesso ai dati in cui sia indicato un profilo di accesso specifico per ogni dato, dettato dai diritti sull'informazione di base, dalle norme o dalle policy in atto, nel rispetto dei principi fondamentali in materia di protezione dei dati personali di cui all'art. 5 del [**GDPR**].

5.2 Richieste di riutilizzo

L'apertura dei dati può essere un'operazione conseguente anche ad una esplicita richiesta da parte di un soggetto interessato. La Direttiva evidenzia che, in questi casi, i tempi di risposta alle

richieste di riutilizzo dei documenti dovrebbero essere ragionevoli ed essere in linea con il tempo necessario per rispondere alle richieste di accesso a un dato documento conformemente ai pertinenti regimi di accesso, come effettivamente recepito con il Decreto, e nel rispetto della normativa unionale e nazionale in materia di protezione dei dati personali (cfr. indicazioni in merito presenti nei paragrafi **4.1** e **5.1.2**).

La Direttiva, inoltre, invita gli Stati membri ad incoraggiare, laddove necessario, la creazione di indici accessibili online per i documenti disponibili in modo da promuovere e agevolare le richieste di riutilizzo. Tale raccomandazione può trovare applicazione nell'ordinamento italiano nelle attività di cui all'art. 53, comma 1-bis del **CAD** volte alla pubblicazione, da parte delle pubbliche amministrazioni, del catalogo dei dati e dei metadati nonché delle relative banche dati in loro possesso, preferibilmente attraverso collegamenti ipertestuali alla sezione del sito in cui sono presenti i relativi dati, le informazioni o i documenti.

Le modalità, i termini e i tempi per le richieste di riutilizzo sono disciplinati dall'art. 5 del Decreto. Sulla base di tali disposizioni e considerato che le pubbliche amministrazioni e gli organismi di diritto pubblico hanno già implementato procedimenti per ottemperare a quanto disposto dall'art. 5 del D. Lgs. n. 33/2013 [**D-LGS-33-2013**] relativo alle richieste di accesso civico, di seguito viene indicata la procedura da seguire per le richieste di riutilizzo di dati e documenti aperti e il trattamento delle stesse.

REQUISITO 18: dlgs36-2006/opendata/req/organization/requests
In caso di richieste di riutilizzo di dati DEVE essere seguita la procedura di esame e trattamento di dette richieste definita nelle presenti Linee Guida, salvo che il titolare dei dati non abbia individuato una propria specifica analoga procedura. Sono tenute comunque a definire termini e modalità di riutilizzo dei dati le imprese pubbliche, gli istituti di istruzione, le organizzazioni che svolgono attività di ricerca, le organizzazioni che finanziano la ricerca, il DIS, l'AISE e l'AISI. In ogni caso, DEVONO essere rispettati i termini temporali previsti dal Decreto.

1. La richiesta di riutilizzo PUÒ essere rivolta ad uno degli uffici indicati all'art. 5, comma 3 del D. Lgs. n. 33/2013 [**D-LGS-33-2013**], ovvero:
    a. all'ufficio che detiene i dati, le informazioni o i documenti;

b. all'Ufficio relazioni con il pubblico;

c. ad altro ufficio indicato dall'amministrazione nella sezione "Amministrazione trasparente" del sito istituzionale;

d. al responsabile della prevenzione della corruzione e della trasparenza, ove l'istanza abbia ad oggetto dati, informazioni o documenti oggetto di pubblicazione obbligatoria ai sensi del D. Lgs. 33/2013 [**D-LGS-33-2013**].

2. Nel caso in cui un'Amministrazione abbia individuato e definito una procedura specifica per la presa in carico e l'espletamento delle richieste di riutilizzo, allora il richiedente DEVE seguire tale procedura by-passando il punto 1. Sulla base del Decreto, sono tenuti a definire specifici termini e modalità di riutilizzo dei dati secondo i rispettivi ordinamenti le imprese pubbliche, gli istituti di istruzione, le organizzazioni che svolgono attività di ricerca, le organizzazioni che finanziano la ricerca, il Dipartimento delle informazioni per la sicurezza (DIS), l'Agenzia informazioni e sicurezza esterna (AISE) e l'Agenzia informazioni e sicurezza interna (AISI).

3. L'esame delle richieste DEVE essere concluso con la formulazione di una **decisione entro 30 giorni**. A supporto di tale attività può essere considerato quanto indicato in relazione all'analisi dei vincoli di cui al par. **5.1.2** o in relazione ai "*Documenti esclusi dall'applicazione*" di cui al par. **1.2**.

4. Nel caso le richieste siano numerose o complesse, allora il termine di cui al punto 3. PUÒ essere **prorogato di ulteriori 20 giorni** previa comunicazione al richiedente entro ventuno giorni dalla richiesta.

5. Se la decisione è positiva, i documenti oggetto della richiesta DEVONO essere **resi disponibili secondo i requisiti definiti nelle presenti Linee Guida**.

6. Se la decisione è negativa, l'ente titolare del dato DEVE **motivare il diniego attraverso un apposito provvedimento** sulla base delle disposizioni del Decreto (per esempio, se i documenti richiesti rientrano tra i documenti esclusi dall'applicazione del Decreto stesso, v. par. **1.2**). Il provvedimento di diniego DEVE includere anche la comunicazione sui mezzi di tutela che il richiedente può esperire sulla base dell'art. 25, commi 4 e 5 della legge 7 agosto 1990, n. 241, mezzi di tutela che per convenienza si riportano al successivo punto 7, fermo restando di fare riferimento in ogni caso ai riferimenti normativi indicati innanzi.

7. In caso di diniego, il richiedente PUÒ presentare ricorso al tribunale amministrativo regionale, ovvero chiedere, nei confronti degli atti delle amministrazioni comunali, provinciali e regionali, al difensore civico competente per ambito territoriale, ove costituito, o a quello competente per l'ambito territoriale immediatamente superiore, che sia riesaminata la suddetta determinazione. Nei confronti degli atti delle amministrazioni centrali e periferiche dello Stato tale richiesta è inoltrata presso la Commissione per l'accesso ai documenti amministrativi istituita presso la Presidenza del Consiglio dei Ministri nonché presso l'amministrazione resistente. Il difensore civico o la Commissione per l'accesso si pronunciano entro trenta giorni dalla presentazione dell'istanza. Scaduto infruttuosamente tale termine, il ricorso si intende respinto. Se il difensore civico o la Commissione per l'accesso ritengono illegittimo il diniego o il differimento, ne informano il richiedente e lo comunicano all'autorità disponente. Se questa non emana il provvedimento confermativo motivato entro trenta giorni dal ricevimento della comunicazione del difensore civico o della Commissione, l'accesso è consentito e quindi l'ente DEVE seguire le indicazioni di cui al punto 6. Qualora il richiedente l'accesso si sia rivolto al difensore civico o alla Commissione, il termine per ricorrere giudizialmente avverso la decisione del difensore civico o della Commissione decorre dalla data di ricevimento, da parte del richiedente, dell'esito dell'istanza presentata al difensore civico o alla Commissione stessa. Se l'accesso è negato o differito per motivi inerenti ai dati personali che si riferiscono a soggetti terzi, la Commissione provvede, sentito il Garante per la protezione dei dati personali, il quale si pronuncia entro il termine di dieci giorni dalla richiesta, decorso inutilmente il quale il parere si intende reso.
8. Nel caso in cui il riutilizzo sia negato perché si tratta di documenti su cui terzi detengono diritti di proprietà intellettuale ai sensi della legge 22 aprile 1941, n. 633, l'ente DEVE indicare la persona fisica o giuridica titolare del diritto, se è nota, oppure il licenziante dal quale il titolare del dato stesso ha ottenuto il materiale. Sono esentati da questa indicazione le biblioteche, comprese quelle universitarie, i musei e gli archivi (cfr. art. 5, comma 4 del Decreto).

5.3 Qualità dei dati

Il miglioramento della qualità dei dati e la maggiore diffusione delle tecniche di misurazione dipendono da vari fattori tra cui l'adesione a modelli di qualità condivisi.

Per determinare la bontà dei dati è necessario definire delle misure attraverso le quali quantificare la qualità dei dati. Lo standard ISO/IEC 25012:2008, divenuto norma italiana UNI CEI ISO/IEC 25012:2014, definisce un insieme di caratteristiche specifiche per la caratterizzazione della qualità dei dati suddivise in "***inerenti***" (accuratezza, aggiornamento (attualità), completezza, consistenza (coerenza), credibilità), "***inerenti e dipendenti dal sistema***" (accessibilità, comprensibilità, conformità, efficienza, precisione, riservatezza, tracciabilità) e "***dipendenti dal sistema***" (disponibilità, portabilità e ripristinabilità).

Sulla base dello standard citato, dal punto di vista inerente, la qualità dei dati si riferisce ai dati stessi, in particolare a: i) valori di dominio dei dati e possibili limitazioni; ii) relazioni di valori di dati; iii) metadati. Riguardo al primo punto, un esempio è rappresentato dalle cosiddette "*regole di business*" definite in relazione allo specifico contesto operativo per verificare una determinata caratteristica. Tali regole implicano che la qualità dei dati sia parte integrante del processo di produzione del dato in modo da semplificare le operazioni che costituiscono il processo stesso applicando tecniche di qualità dei dati alle informazioni già durante l'elaborazione.

Con riferimento alla qualità dei dati dipendente dal sistema, detta qualità dipende dal dominio tecnologico in cui i dati sono utilizzati.

La qualità dei dati, sebbene riferita a quelli riportati nei siti istituzionali nel rispetto degli obblighi di pubblicazione previsti dalla legge, è anche un obbligo derivante dall'art. 6 del decreto legislativo n. 33/2013 [**D-LGS-33-2013**] secondo cui si deve assicurare "*l'integrità, il costante aggiornamento, la completezza, la tempestività, la semplicità di consultazione, la comprensibilità, l'omogeneità, la facile accessibilità, nonché la conformità ai documenti originali in possesso dell'amministrazione, l'indicazione della loro provenienza e la riutilizzabilità […]*".

Lo stesso art. 6, al comma 2, precisa che "*l'esigenza di assicurare adeguata qualità delle informazioni diffuse non può, in ogni caso, costituire motivo per l'omessa o ritardata pubblicazione dei dati, delle informazioni e dei documenti*". Questo implica che sia possibile pubblicare dati e documenti incompleti relativamente alle caratteristiche di qualità indicate per rispettare i tempi di pubblicazione previsti dalle norme. SI RACCOMANDA, però, di indicare i motivi del mancato adempimento derivante dall'art. 6, comma 1 e di procedere successivamente all'aggiornamento dei dati garantendo il rispetto delle caratteristiche di qualità richiamate.

La Determinazione Commissariale n. 68/2013 di AgID[41], relativa alle regole tecniche per l'identificazione delle basi di dati critiche tra quelle di interesse nazionale specificate sulla base dell'art. 60 del **CAD**, disponeva che venisse garantito il rispetto di quattro caratteristiche, delle quindici previste dallo Standard ISO/IEC 25012, ovvero:

- **accuratezza** (sintattica e semantica) - il dato e i suoi attributi rappresentano correttamente il valore reale del concetto o dell'evento a cui ci si riferiscono;
- **coerenza** - il dato e i suoi attributi non presentano contraddittorietà rispetto ad altri dati del contesto d'uso dell'amministrazione che detiene il dato;
- **completezza** – il dato risulta esaustivo, sia per tutti i suoi valori attesi e sia rispetto alle entità relative (fonti) che concorrono alla definizione del procedimento a cui si riferisce;
- **attualità** (o tempestività di aggiornamento) - il dato e i suoi attributi sono del "giusto tempo" (sono aggiornati) rispetto al procedimento a cui si riferiscono.

Raccomandazione 8: dlgs36-2006/opendata/req/organization/quality
SI RACCOMANDA di garantire, per tutti i dati in generale e per quelli resi disponibili per il riutilizzo, in particolare, il rispetto almeno delle quattro caratteristiche di qualità dei dati (ovvero accuratezza, coerenza, completezza e attualità), fra le 15 previste dallo Standard ISO/IEC 25012 come da indicazioni della Determinazione Commissariale n. 68/2013 di AgID.
Per la misura delle suddette caratteristiche, fare riferimento allo Standard ISO/IEC 25024.

Altre caratteristiche di qualità sono garantite per via degli adempimenti richiesti da specifiche norme di legge o regole tecniche. Per esempio:

- l'**accessibilità** è un obbligo derivante dalla legge n.4/2004 e dalle relative Linee Guida AgID;
- la **riservatezza** è correlata alle indicazioni derivanti dal [**GDPR**];
- la **disponibilità** include il disaster recovery;
- la **ripristinabilità** può essere fornita attraverso meccanismi di backup.

[41] https://www.agid.gov.it/sites/default/files/repository_files/circolari/dt_cs_n.68_-_2013dig_-regole_tecniche_basi_dati_critiche_art_2bis_dl_179-2012_sito.pdf

Il passo successivo è quantificare le caratteristiche di qualità in termini di misure, individuando delle soglie che consentano di discriminare la bontà o meno di un dato rispetto alla caratteristica in esame. La fase di valutazione della qualità dei dati è importante in tutti i sistemi informativi indipendentemente dalla scelta/necessità di procedere alla loro apertura. Con l'adozione di politiche di apertura dei dati, la qualità dei dati assume un ruolo ancora più rilevante in quanto elemento per la certificazione della bontà dei dati forniti e soprattutto dell'appropriatezza rispetto all'utilizzo che del dato si vuole fare.

Lo Standard UNI CEI ISO/IEC 25024:2016 "Misurazione della qualità del dato" estende l'UNI CEI ISO/IEC 25012 al campo delle misurazioni, definendo 63 misure di qualità applicabili alle 15 caratteristiche di qualità dei dati, con le relative funzioni di calcolo.

Secondo lo standard, la qualità va ricercata durante l'intero ciclo di vita del dato, un esempio del quale, riportato nello standard stesso, è composto dalle fasi di: progettazione, acquisizione, integrazione con altri dati, elaborazione, memorizzazione, uso, cancellazione.

Nei metadati è possibile indicare qual è il livello di qualità dei dati specificando sia il risultato delle misure applicate che fornendo informazioni relative alle caratteristiche di qualità (per esempio, la frequenza di aggiornamento relativamente all'attualità, l'associazione di una impronta crittografica basata su funzioni di hash relativamente alla integrità dei dati, …).

Nella tabella che segue si riporta un insieme esemplificativo di misure, sulle 24 definite nello standard ISO per le stesse caratteristiche, a supporto delle attività di valutazione della qualità dei dati delle amministrazioni.

Caratteristiche	Descrizione	Misure e funzioni di misura
Completezza	il grado per cui il dato associato a un'entità presenta valori per tutti gli attributi attesi e relative istanze in un certo contesto.	Si individuano i seguenti livelli di completezza: 1. completezza di schema: percentuale di valori nulli per concetti e proprietà rispetto al numero totale di valori attesi; 2. completezza dei record: numero di dati elementari associati a un valore non nullo in un record, rispetto al numero di dati elementari del record

		per cui può essere misurata la completezza; 3. completezza di popolazione: percentuale di valori nulli rispetto a una popolazione di riferimento. Si noti che non sempre valori mancanti indicano incompletezza. Per esempio: si supponga di considerare dati relativi ai musei italiani e ai loro canali di contatto (telefono ed email). Può capitare che i musei abbiano tutti un indirizzo email ma non per tutti è presente un numero di telefono.
Accuratezza	Il grado in cui gli attributi rappresentano in maniera corretta il valore reale del dato in uno specifico contesto	Si individuano due tipi di accuratezza: 1. sintattica: ad esempio Merio invece che Mario 2. semantica: ad esempio nel caso in cui si utilizzi Marco Rossi intendendo invece un'altra persona per es., Mario Rossi Una misura dell'accuratezza è data dal rapporto tra gli attributi dei dati che hanno valori accurati sintatticamente/semanticamente sul numero di attributi dei dati per i quali è richiesta accuratezza sintattica/semantica.
Coerenza	Il grado in cui gli attributi del dato non sono in contraddizione con altri dati in uno specifico contesto	Per poter valutare la coerenza una misura è quella che consente di identificare le violazioni di regole semantiche definite su alcuni elementi dei dati.

		Per esempio, se una persona è "patentata" non può essere possibile che la sua età sia "17 anni". Essa può essere calcolata come il rapporto tra il numero di attributi dei dati i cui valori sono semanticamente corretti nel dataset sul numero di attributi dei dati per i quali sono state definite delle regole semantiche. Altra misura consiste nel rapporto tra il numero di valori duplicati per ogni attributo della base dati e il numero totale degli elementi della base dati.
Tempestività	Il grado in cui gli attributi del dato sono al "giusto tempo" rispetto al contesto di riferimento	La metrica è basata sull'uso dei metadati che indicano quando il dato è stato aggiornato l'ultima volta. Sulla base di questi metadati, si distinguono poi: 1. dati con periodicità di aggiornamento nota: in questo caso è possibile calcolare la tempestività in maniera esatta identificando se la data di ultima modifica del dato rispetto al tempo di misurazione ricada nell'intervallo della frequenza di aggiornamento; 2. dati con periodicità di aggiornamento media: in questo caso è possibile calcolare la tempestività media con una percentuale di errore.

Tabella 3 - Caratteristiche di qualità e relative misure

Lo Standard ISO/IEC 25012 è applicabile a tutte le tipologie di dati. Nel caso di dati territoriali, uno standard specifico di riferimento per la qualità è l'ISO 19157 "Geographic information -- Data quality". Alcuni elementi e misure di tale Standard sono utilizzati per

identificare requisiti e raccomandazioni per la qualità dei dati nella sezione "7 - Data quality" delle specifiche sui dati definite per ciascun tema INSPIRE[42]. Per i dati territoriali di cui ai temi INSPIRE, pertanto, dovranno essere considerati gli elementi e le misure definite nelle specifiche citate innanzi, nonché i sistemi di validazione e controllo nazionali ed europei. Per i dati territoriali che fanno riferimento a specifiche diverse, dovranno essere adottati sistemi di verifiche di qualità a cura di soggetti qualificati diversi da quelli che hanno prodotto il dato.

[42] https://inspire.ec.europa.eu/Technical-Guidelines/Data-Specifications/2892

Capitolo 6
Aspetti legali e di costo

Gli articoli 7 e 8 del Decreto prevedono disposizioni specifiche su tariffazione e licenze standard da adottare per il riutilizzo dei dati. Atteso che il principio preminente stabilito da Direttiva e Decreto è che il riutilizzo dei documenti non dovrebbe essere soggetto a condizioni, nel presente capitolo vengono fornite le indicazioni utili per supportare i soggetti titolari del trattamento dei dati nella scelta della licenza da applicare ai dati aperti e nell'eventuale addebito, a carico del riutilizzatore, dei costi marginali sostenuti effettivamente per la riproduzione, la messa a disposizione e la divulgazione dei dati, nonché per l'anonimizzazione dei dati personali o per le misure adottate per proteggere le informazioni commerciali a carattere riservato. In determinati casi è possibile anche determinare tariffe superiori ai costi marginali.

Altre indicazioni riguardano il principio di non discriminazione e gli accordi di esclusiva tra enti pubblici e partner privati da evitare, per quanto possibile, ma che, in alcuni casi, possono risultare necessari.

6.1 Licenze e condizioni di riutilizzo

Una delle caratteristiche dei dati di tipo aperto - come previsto nella definizione data all'art. 1, comma 1, lettera l-ter) del [**CAD**] - è quella di essere "***disponibili secondo i termini di una licenza o di una previsione normativa che ne permetta l'utilizzo da parte di chiunque, anche per finalità commerciali, in formato disaggregato***".

Principio preminente indicato dal Decreto è che il riutilizzo dei documenti non debba essere soggetto a condizioni ("*Il riutilizzo di documenti non è soggetto a condizioni, [...]*", cfr. art. 8, comma 2); in coerenza con tale indicazione, anche nel caso di richieste di riutilizzo (cfr. art. 5, comma 2 del Decreto), è previsto che "*in caso di decisione positiva, i documenti sono resi disponibili, ove possibile, in forma elettronica e, se necessario, attraverso una licenza*", sottolineando in tal modo, indirettamente, la possibilità di non apporre alcuna restrizione (quindi, alcuna licenza) al riutilizzo dei dati.

Tuttavia, la Direttiva precisa che in alcuni casi giustificati da un obiettivo di pubblico interesse, PUÒ essere utilizzata una licenza che impone al titolare del dato condizioni di riutilizzo riguardanti "*questioni quali la responsabilità, la protezione dei dati di carattere personale, l'uso corretto dei documenti, la garanzia di non alterazione e la citazione della fonte*" (cfr. Considerando 44).

In quest'ultimo caso (applicazione di specifiche condizioni), il Decreto stabilisce che tali condizioni debbano essere oggettive, proporzionate e non discriminatorie, nonché giustificate da un pubblico interesse (cfr. art. 8, comma 2).

In tale contesto, l'apposizione di una licenza, oltre a identificare e "definire" correttamente i dati aperti, costituisce uno strumento funzionale a garantire certezza circa l'effettiva riutilizzabilità dei dati; certezza che costituisce un presupposto essenziale alla valorizzazione dell'informazione, specie nel settore pubblico. Seppure, quindi, in assenza di specifica licenza operi il principio dell'"*open by default*" previsto dall'art. 52 del [**CAD**], SI RACCOMANDA di apporre sempre una licenza ai dataset pubblicati, in modalità tali da renderla facilmente individuabile e comprensibile.

Naturalmente, restano esclusi dall'ambito di applicazione del presente paragrafo, così come della sopra menzionata raccomandazione, i testi degli atti ufficiali dello Stato e delle Amministrazioni pubbliche, sia italiane che straniere, che ricadono direttamente in pubblico dominio ai sensi dell'art. 5 della legge n. 633/1941 sul diritto d'autore.

Nel contesto sopra descritto, in particolare, il Decreto dispone l'**utilizzo di licenze standard disponibili in formato digitale** ("*Le pubbliche amministrazioni (...) adottano licenze standard, disponibili in formato digitale, per il riutilizzo dei propri documenti*", cfr. art. 8, comma 1). Tali licenze standard DEVONO comunque prevedere il minor numero possibile di restrizioni al riutilizzo (limitando, per esempio, le restrizioni alla sola indicazione della fonte), che consentano, pertanto, a chiunque di accedere liberamente a dati e contenuti, nonché di utilizzarli, modificarli e condividerli liberamente e per qualsiasi finalità, fermo restando il rispetto delle regole in materia di protezione dei dati personali (cfr. par. **4.1** e par. **5.1.2**).

Raccomandazione 9: dlgs36-2006/opendata/rec/licenses/attribution

SI RACCOMANDA di restringere le condizioni di cui alla licenza apposta ai dati alla sola attribuzione, fatta eccezione per le regole in materia di riutilizzo dei dati personali (cfr. par. 4.1 e par. 5.1.2).

Ricordiamo, inoltre, che, in linea con il principio che vuole l'apposizione del minor numero possibile di restrizioni, l'art. 7 del D. Lgs. 33/2013 [**D-LGS-33-2013**], con riferimento a documenti, informazioni e dati oggetto di pubblicazione obbligatoria, dispone che siano "*pubblicati in formato di tipo aperto ai sensi dell'art. 68*[43] *del Codice dell'amministrazione digitale, di cui al decreto legislativo 7 marzo 2005, n. 82, e sono riutilizzabili ai sensi del decreto legislativo 24 gennaio 2006, n. 36, del decreto legislativo 7*

[43] Si evidenzia che la definizione di formato aperto non è più contenuta all'interno del citato art. 68 del D. Lgs. 82/2005, bensì all'art. 1, comma 1, lett. l-bis del medesimo decreto.

marzo 2005, n. 82, e del decreto legislativo 30 giugno 2003, n. 196, senza ulteriori restrizioni ***diverse dall'obbligo di citare la fonte*** *e di rispettarne l'integrità*".

Tale previsione è coerente con l'impostazione sopra richiamata, che – fatta eccezione per le regole in materia di riutilizzo dei dati personali (cfr. par. **4.1** e par. **5.1.2**) – vede nella "attribuzione" l'unica condizione liberamente e sostanzialmente apponibile ai dati.

SI RACCOMANDA, inoltre, di valutare sempre, nel processo di apertura, la conformità ad ulteriori discipline che impattino sull'apertura dei dati, in primis la normativa unionale e nazionale in materia di protezione dei dati personali come già richiamato nei paragrafi precedenti, ricordando che l'art. 4 del Decreto fa salva tale specifica disciplina nell'utilizzo e riutilizzo di dati.

L'utilizzo di licenze standard favorisce, infine, l'immediata comprensibilità delle stesse e l'uniformità nel loro riutilizzo.

Considerato quanto sopra esposto, si ritiene, quindi, che l'applicazione di condizioni ulteriori rispetto all'attribuzione (quale, ad esempio, la cd. SA - "share alike" - condivisione) presenti aspetti di potenziale criticità, oltre che pratica nel caso di uso di fonti diverse diversamente licenziate, anche rispetto alla compatibilità con l'attuale normativa.

Raccomandazione 10: dlgs36-2006/opendata/rec/conditions/more-conditions

SI RACCOMANDA di limitare l'uso di licenze con condizioni ulteriori rispetto alla sola attribuzione solo ai casi strettamente necessari.

Rispetto alla specifica licenza da apporre, si rammenta che ad oggi, si utilizzano numerose licenze standard, che possono essere suddivise in tre gruppi:

- **Licenze di sola attribuzione**
  - **CC-BY:** prodotta dall'omonimo movimento internazionale (www.creativecommons.org) in diverse versioni successive; nella versione attuale (4.0), consente al licenziatario di condividere e modificare, per qualsiasi finalità, con la sola restrizione dell'attribuzione al licenziante. A differenza di precedenti versioni, le condizioni si applicano anche con riferimento ai diritti "sui generis" e l'*attribution* implica il richiamo di fonte, copyright, ecc. nella misura richiamata dal licenziante e può essere assolta in ogni forma "ragionevole". Vieta inoltre l'apposizione di restrizioni ulteriori, anche di natura tecnologica e richiede indicazione delle modifiche;

- **CDLA-permissiva 1.0**: consente al licenziatario di condividere e modificare, con la sola attribuzione al licenziante e citazione della licenza. Incoraggia l'arricchimento e il miglioramento dei dati e la produzione di opere derivate/mashup, senza creare vincoli con i dati di provenienza. Non impone obblighi o restrizioni ai dati "migliorati" (derivati e/o di mashup) e contiene il concetto di "risultato" - non condizionato - proprio delle elaborazioni algoritmiche;
- **IODL 2.0**: consente al licenziatario di condividere e modificare, per qualsiasi finalità, con la sola restrizione dell'attribuzione al licenziante, comprensiva del nome del soggetto che fornisce il dato, includendo, se possibile, il link alla licenza. Contiene riferimento alla normativa nazionale sul diritto d'autore e sui dati personali;
- **ODC-BY**: consente al licenziatario di condividere e modificare, per qualsiasi finalità, con la sola restrizione dell'attribuzione al licenziante. Prevede il diritto sui-generis, ma precisa espressamente che non regola anche i contenuti della banca dati;

o **Licenze di Attribuzione e Condivisione**:

- **CC-BY-SA**: consente al licenziatario di condividere e modificare, per qualsiasi finalità, con la restrizione dell'attribuzione al licenziante, con la duplice restrizione dell'attribuzione al licenziante e della redistribuzione del prodotto derivato con la stessa licenza dell'originale (o versione successiva). Vieta l'apposizione di restrizioni ulteriori, anche di natura tecnologica;
- **CDLA - Condivisione 1.0**: consente al licenziatario di utilizzare e pubblicare i dati per il riutilizzo, con la duplice restrizione dell'attribuzione al licenziante e della pubblicazione con la stessa licenza. Incoraggia l'arricchimento e il miglioramento dei dati e la produzione di opere derivate/mashup, senza creare vincoli con i dati di provenienza. Impone ai dati "migliorati" (derivati e/o di mashup) l'uso della stessa licenza, ma conserva la libertà d'uso incondizionata dei "risultati";
- **IODL 1.0**: consente al licenziatario di condividere e modificare, per qualsiasi finalità, con la duplice restrizione dell'attribuzione al licenziante, comprensiva

del nome del soggetto che fornisce il dato, includendo, se possibile, il link alla licenza, e della condivisione del prodotto derivato o di mashup con la stessa licenza;

- **OdBl**: specifica per i database, consente al licenziatario di utilizzare, condividere, modificare, integrare e redistribuire il database, con la duplice restrizione dell'attribuzione al licenziante (anche per i prodotti derivati) e dell'uso della stessa licenza. Contempla il concetto di "produced work", ovvero di elaborato dal db ma diverso da quest'ultimo, che può essere diversamente licenziato (salva citazione fonte). Consente l'apposizione di restrizioni ulteriori, anche di natura tecnologica, a condizione che almeno una copia rimanga sempre libera;

- **Waiwer**
  - **CC0**: come noto, non è una vera e propria licenza, ma una rinuncia preventiva all'esercizio dei diritti in qualsiasi modo previsti o connessi al diritto d'autore. Si parla, a riguardo, di attribuzione (o donazione) al pubblico dominio.

Oltre a quelle citate, va ricordata la **CDLA 2.0 permissive**, che si pone al limite del waiwer, posto che richiede, di fatto, il solo richiamo del testo della licenza (oltre a richiamare la nozione già citata di "risultati").

Tutte le ulteriori licenze Creative Commons sono classificabili come:

- licenze che **NON consentono opere derivate**;

oppure come

- licenze che **NON consentono l'uso commerciale**.

REQUISITO 19: dlgs36-2006/opendata/req/conditions/nd-nc-licenses

I titolari dei dati disponibili per il riutilizzo NON DEVONO:

- utilizzare licenze che non consentano opere derivate o uso commerciale;
- utilizzare licenze di tipo proprietario.

Le licenze suindicate sono raffigurate nella Figura seguente.
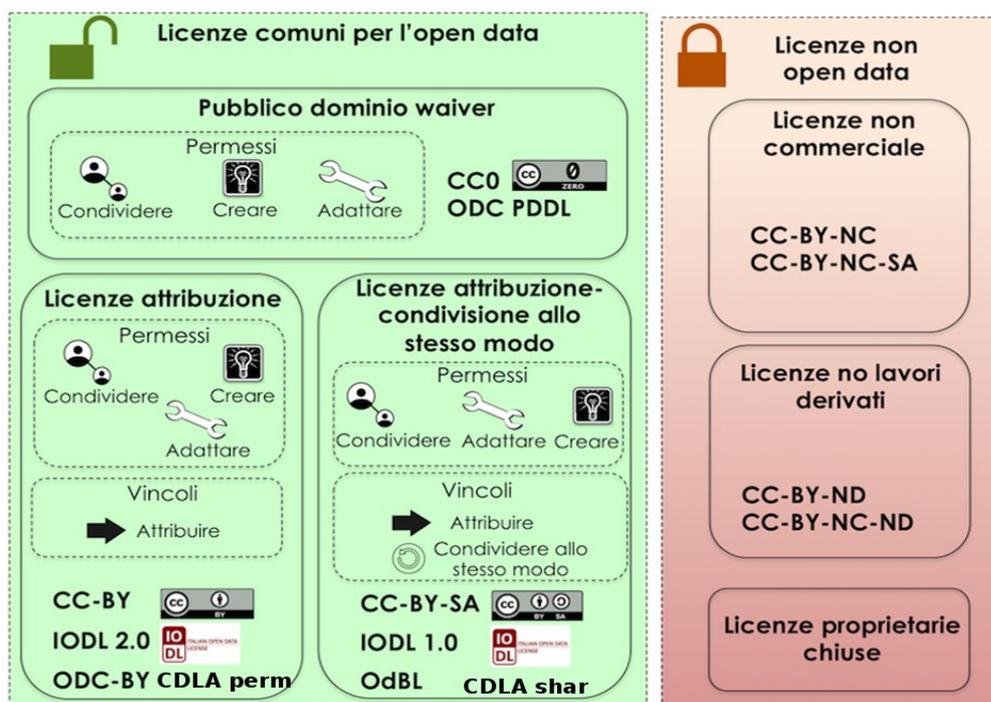
Figura 3 - Licenze

Alla luce del Considerando (44) della Direttiva, le presenti Linee Guida valutano come ragionevole motivo di pubblico interesse l'adozione di una licenza standard omogenea, funzionale a preservare l'interesse parimenti fondato di conservare traccia della fonte "pubblica" del dato e, in particolare, per questioni attinenti all'affidabilità dello stesso (a tutela, peraltro, anche del riutilizzatore), valutano come, in generale, l'unica condizione ammissibile sia la "attribuzione".

> **Raccomandazione 11: dlgs36-2006/opendata/rec/conditions/sa**
> SI RACCOMANDA di limitare l'uso della clausola di "condivisione" ("share-alike" - SA) solo ai casi in cui sia motivatamente necessaria ovvero previa verifica di impossibilità di rilascio con licenza CC BY 4.0, ad esempio, in ragione dell'uso non altrimenti gestibile di una fonte già rilasciata con licenza SA.

Particolare cautela, rispetto ai database, va utilizzata anche nella scelta della licenza standard, nonché nella gestione dell'"attribuzione", in quanto le relative condizioni sono spesso declinate in modo differente nelle differenti licenze. Inoltre, non sempre tutte le licenze standard presentano condizioni riferite ad un bene come una "base di dati" (si fa qui riferimento soprattutto alle versioni delle Creative Commons precedenti alla 4.0, a quanto consta ancora utilizzate: vedasi per i dettagli in allegato la nota espositiva delle principali licenze).

Si ricorda, a riguardo, che le basi di dati godono di una duplice tutela ai sensi della normativa nazionale e comunitaria, ovvero la tutela quale opera creativa, ove ne ricorrano i presupposti, e la tutela del cosiddetto "diritto sui generis", che tutela lo sforzo di costituzione di una banca dati, anche magari non creativa, da una estrazione totale o sostanziale, separando la tutela della base dati in quanto tale dalla tutela eventuale dei singoli contenuti.

Raccomandazione 12: dlgs36-2006/opendata/rec/conditions/cc

SI RACCOMANDA di non utilizzare le licenze Creative Commons precedenti alla 4.0, in cui i diritti sui generis non erano citati/previsti (2.5) o erano richiamati come meramente rinunciati (3.0).

Raccomandazione 13: dlgs36-2006/opendata/rec/conditions/iodl

SI RACCOMANDA di evitare quelle licenze che – per quanto ben impostate – presentano forti caratteristiche di localizzazione, anch'esse potenzialmente costituenti elementi di ambiguità in caso di riutilizzo e mashup (come la IODL).

In relazione a quanto sopra riportato, tenuto conto del contesto normativo di riferimento e delle indicazioni in tema di licenze contenute nella Comunicazione della Commissione 2014/C - 240/01, è necessario, almeno per i dati aperti "nativi" - riconducibili essenzialmente ai dati che vengono prodotti dalle amministrazioni pubbliche e dagli altri soggetti destinatari delle presenti Linee Guida, nell'adempimento delle proprie funzioni istituzionali - **fare riferimento ad una licenza unica aperta, che garantisca la libertà di riutilizzo, che sia internazionalmente riconosciuta e che consenta di attribuire la paternità dei dataset (attribuire la fonte).** In alternativa, potrà essere adottata la CC0. Salvo il caso di diversa espressa disposizione, peraltro, il requisito relativo alla "attribution" si intende soddisfatto mediante citazione, nella forma più appropriata al mezzo utilizzato, dell'ente titolare del dato riutilizzato.

REQUISITO 20: dlgs36-2006/opendata/req/conditions/ccby4

Per i nuovi dati aperti nativi, salvo quanto precisato nel **REQUISITO 21**, DEVE essere applicata la licenza CC-BY nell'ultima versione disponibile (al momento della stesura delle presenti Linee Guida, la 4.0), presupponendo altresì l'attribuzione automatica di tale licenza nel caso di applicazione del principio "open data by default", di cui all'articolo 52 del **CAD**.

REQUISITO 21: dlgs36-2006/opendata/req/conditions/no-ccby4
L'adozione, qualora possibile e/o previsto, di una licenza diversa dalla CC-BY 4.0 o CC0 DEVE essere formalmente motivata, anche alla luce dei principi espressi dalla Direttiva, salvo che sia stata adottata una licenza altrettanto compatibile come la CDLA 2.0 permissive ovvero qualsiasi altra licenza aperta equivalente o meno restrittiva, che consenta il riutilizzo salvo obbligo di attribuzione, dando credito al concedente.
Per le serie di dati di elevato valore vale quanto indicato nel Regolamento (UE) di esecuzione n. 2023/138.

Raccomandazione 14: dlgs36-2006/opendata/rec/conditions/update
SI RACCOMANDA ai titolari che hanno già pubblicato set di dati con licenze diverse da quelle sopra richiamate, incluse versioni della CC-BY precedente alla 4.0, di valutare il rinnovo della licenza, adeguandola alle indicazioni suddette, individuando nel caso le ragioni eventualmente impedienti tale aggiornamento.

6.1.1 Compatibilità tra licenze

Anche le licenze "aperte", analogamente alle licenze open source, presentano differenti gradi di apertura (non sono, quindi, sempre "aperte allo stesso modo"), ovvero prevedono condizioni che, pur autorizzando il riutilizzo, possono non rendere percorribile un riutilizzo "mescolato" tra più fonti; tali condizioni, pertanto, possono non consentire una successiva pubblicazione/un utilizzo nel rispetto di tutte le condizioni previste da ciascuna licenza (incompatibilità).

Per fare un esempio, due licenze aperte cd. "share alike" – che richiedono di rilasciare ogni evoluzione successiva con la medesima licenza nei medesimi termini – permettono il rilascio di un mashup solo se tra loro identiche, o se tra loro sia stata riconosciuta una eventuale equivalenza (v. infra). Inoltre, anche nel caso di licenze fra loro compatibili, si segnala che dovranno sempre essere rispettate le relative condizioni di ridistribuzione, tenendo, altresì, conto delle eventuali diversità sull'ambito di applicazione e/o di esenzione: per esempio, alcune licenze richiedono di segnalare le modifiche, o escludono dal perimetro di applicazione le elaborazioni algoritmiche o i prodotti derivati di natura diversa dal database di origine, etc... Anche a tale fine, quindi, si è di seguito provato a evidenziare le principali condizioni e/o peculiarità presenti nelle principali licenze standard, individuando in quali trovino applicazione (v. **Tabella 4**).

Cod.	Condizioni da osservare / specifiche di applicazione	Licenze impattate						
A	segnalazione modifica cambiamenti	CDLA 1 perm	CDLA 1 SHAR		CC-BY			
B	testo licenza: con riferimento al dataset originale, riportare il testo della licenza e/o inserire link (tra [ ] per la IODL in quanto precisa "se possibile")	CDLA 1 perm	CDLA 1 SHAR	CDLA 2 perm	CC-BY	[IODL 2.0]	IODL 1.0	OdBL
C	attribution: rispetto specifiche condizioni	CDLA 1 perm	CDLA 1 SHAR		CC-BY	IODL 2.0	IODL 1.0	OdBL
D	output ulteriore realizzato ["Produced work" - es. Mappa da db geografico]; richiede sola attribution							OdBL
E	risultato da "computational use" senza condizioni	CDLA 1 perm	CDLA 1 SHAR / perm 1 e 2	CDLA 2 perm				
F	prevede la nozione di "collective database"							OdBL
G	limiti DRM				CC-BY			OdBL

Tabella 4 - Condizioni da osservare / specifiche di applicazione

Sulla base di queste condizioni e al fine di fornire un supporto di riferimento si propone di seguito una valutazione di compatibilità relativa:

- evoluzione di una precedente singola fonte (v. **Tabella 5**);
- creazione di un nuovo dataset costituito da più fonti diverse e diversamente licenziate (v. **Tabella 6**)

rispetto a cui si è provato a distinguere:

- per la **Tabella 5**, le evoluzioni:
  - rilasciabili con la licenza indicata nella riga propria dell'opera derivata (caselle verdi e simbolo);
  - da valutare (caselle gialle e simbolo);
  - non rilasciabili (caselle rosse e simbolo);
- per la **Tabella 6**, le combinazioni:
  - rilasciabili con una delle licenze indicate nelle presenti linee guida (caselle verdi e simbolo);
  - rilasciabili, ma con licenze differenti e sconsigliate e/o soggette a particolari punti di attenzione (caselle gialle e simbolo);
  - non rilasciabili (caselle rosse e simbolo).

Le relative matrici di compatibilità costituiscono il risultato di una valutazione interpretativa del Gruppo di Lavoro nel tentativo di fornire un quadro complessivo il più articolato e completo possibile. Si evidenzia, pertanto, che gli schemi così forniti costituiscono una semplice guida per l'interprete che non può costituire "interpretazione autentica" della compatibilità delle licenze stesse, in assenza di una concorde presa di posizione da parte dei soggetti che le hanno

prodotte. È quindi raccomandata sempre una verifica caso per caso, specie in relazione alle ipotesi segnalate come "dubbie" (v. legende, colore giallo).

a) evoluzione di una precedente singola fonte (opera "derivata")

In questo caso, la licenziabilità della soluzione e le relative condizioni saranno influenzate solo dalla licenza originaria: la tabella che segue è relativa alla possibilità di produrre un dataset "derivato", in linea con le indicazioni fornite con le presenti Linee Guida ovvero, in CC-BY 4.0 o, in subordine, CC0 o CDLA 2.0 permissive (v. **Tabella 5**).

licenza opera "derivata" / licenza opera originaria	**CC0**	**CC-BY 4.0**	**CDLA 2.0 perm**	**CC-BY-SA 4.0**	**ODbL**
CC0	👍	👍 [A-B-C]	👍	👍	👍
CC-BY 4.0	👎	[A-B-C]	⚠ *[A-B-C] **	👍 [A-B-C]	⚠ *[A-B-C]
CDLA 2.0 perm	⚠ [A] ***	👍 [B-E]	👍 [B-E]	👍 ** [B-E]	👍 [B-E]
IODL 2.0	👎	👍 [B-C]	⚠ [B-C] **	👍 [B-C]	👍 [B-C]
CC-BY-SA 4.0	👎	👎	👎	👍 [A-B-C]	👎
ODbL	👎	👎	👎	👎	👍 [(A) B-C-D-F-G]
CDLA 1.0 shar	👎	👎	👎	👎	👎

Tabella 5 - Licenze applicabili all'opera derivata in funzione della licenza originaria

Legenda Tabella 5:

👍	pubblicabile con licenza indicata come opera "derivata"
👎	non ripubblicabile
⚠	compatibilità espressamente messa in discussione[44][45] e/o potenzialmente discutibile in base a DRM e/o modalità di attribuzione OPPURE compatibilità richiamata da documentazione relativa solo a una delle licenze (es. IODL 1.0 rispetto a CC BY SA)[46]

* verificare gestibilità *attribution* e DRM

[44] v. COMMISSION DECISION of 22.2.2019 adopting Creative Commons as an open licence under the European Commission's reuse policy
[45] v. https://blog.openstreetmap.org/2017/03/17/use-of-cc-by-data/
[46] https://it.wikipedia.org/wiki/Italian_Open_Data_License

** il sito della CDLA (v. box infra) ritiene compatibile la CC BY 4.0 con il rilascio in CDLA 2.0 permissive, a condizione di rispettare l'*attribution* originale (anche se non raccomandato, per non creare "strati" di attribuzione", come da wiki delle CC di cui al link nel box "Risorse utili", "Adapter's license chart"). Analogo ragionamento a fortiori parrebbe essere applicabile per la IODL 2.0. Sempre come esempio, se si elabora un db rilasciato originariamente in CC BY SA, o in OdBL, sì dovrà rilasciare anche il nuovo DB in CC BY SA o rispettivamente OdBL (salvo l'eccezione del "produced work" per l'OdBL)

*** il sito della CDLA dichiara compatibile il rilascio di un'opera derivata sotto CC0 partendo da un'opera originaria sub CDLA 2.0 permissive a condizione di riportare il testo di quest'ultima (che tuttavia parrebbe condizione aggiuntiva alla CC0).

b) creazione di un nuovo dataset costituito da più fonti diverse e diversamente licenziate

In questo caso, è necessario verificare che le licenze originarie non risultino incompatibili con la pubblicazione (v. **Tabella 6**) relative alle peculiarità del caso.

licenza opera 1 / licenza opera 2	CC0	CC-BY 4.0	CDLA 2.0 perm	IODL 2.0	CC-BY-SA 4.0	ODbL	CDLA 1.0 shar
CC0				IODL 2.0	CC BY SA	ODbL	CDLA1shar
CC-BY 4.0			CC-BY 4.0 [riportando testo CDLA] o CDLA [conservando *attribution* C CC BY] *		CC BY SA	*[D] OdBL	* CDLA1shar
CDLA 2.0 perm		CC-BY 4.0 [riportando testo CDLA] o CDLA [conservando *attribution* C CC BY] *		Entrambe compatibili con CC BY 4.0 [B – E CDLA]	CC BY SA [B – E CDLA]	OdBL [B – E CDLA]	CDLA1 shar [B – E CDLA]
IODL 2.0			Entrambe compatibili con CC BY 4.0 [B – E CDLA]		CC BY SA	OdBL	CDLA1 shar
CC-BY-SA 4.0	CC BY SA	CC BY SA	CC BY SA [B e E CDLA]	CC BY SA	CC BY SA		
ODbL	OdBL	* ODbL [D]	OdBL [B – E CDLA]	OdBL		OdBL	
CDLA 1.0 shar	CDLA1 shar	* CDLA1shar	CDLA1shar [B – E CDLA]	CDLA1 shar			CDLA1shar

Tabella 6 - Matrice di compatibilità tra licenze

Legenda Tabella 6:

👍	pubblicabile con licenza CC BY 4.0, CC0 o CDLA permissive
👎	non ripubblicabile
⚠	pubblicabile, ma con licenza diversa da CC BY / CC0 / CDLA permissive e/o compatibilità espressamente messa in discussione[47] [48] e/o potenzialmente discutibile in base a DRM e/o modalità di attribuzione OPPURE compatibilità richiamata da documentazione relativa solo a una delle licenze (es. IODL 1.0 rispetto a CC BY SA)[49]

* verificare gestibilità *attribution* e DRM

Nota Bene:

- la combinazione tra CC0 e altra licenza produce sempre un'opera rilasciabile ai sensi della licenza più restrittiva;
- la combinazione tra IODL 2.0 e CDLA 2.0 permissive viene proposta in CC BY 4.0 in quanto entrambe le licenze sono dichiarate dall'autore come compatibili con essa.

Quest'ultima tabella espone, quindi, una matrice di compatibilità - necessariamente limitata alle principali licenze standard - che distingue i casi in cui la combinazione delle fonti permetta di licenziare la soluzione come da indicazioni (verde), comunque aperta (giallo) o produca un blocco (rosso).

In particolare, si segnala che all'interno di ogni singola casella, si sono richiamate le condizioni (con la relativa lettera di cui alla **Tabella 4**) – comunque da rispettare nel rilascio dell'opera derivata o combinata (ad esempio, nel rilasciare un'opera derivata con CC BY SA 4.0 partendo da un'opera sub CC BY 4.0, dovrà essere rispettata l'attribuzione dell'opera originale, etc.).

In entrambe le tabelle 5 e 6 (come da legenda in calce) si è provato a dettagliare meglio alcuni aspetti che, pur non impedendo il riutilizzo, sono da considerarsi punti di attenzione.

Si fa riferimento, a titolo di esempio:

- al rischio di cumulo delle attribuzioni (cd. "stack of attribution"), presente anche nelle licenze solo "permissive" (mera attribuzione), ma diverse tra loro, le quali

[47] v. COMMISSION DECISION of 22.2.2019 adopting Creative Commons as an open licence under the European Commission's reuse policy

[48] v. https://blog.openstreetmap.org/2017/03/17/use-of-cc-by-data/

[49] https://it.wikipedia.org/wiki/Italian_Open_Data_License

presentano, a volte, specifiche diverse, in merito alle modalità con cui assolvere all'obbligo di "attribution". Tali specifiche possono rendere complessa la gestione delle licenze (si pensi, per esempio, ai dati geografici, per i quali l'uso di fonti diverse e ricorsive può rendere difficile detto governo);

- ai limiti all'apponibilità di misure tecnologiche di protezione, in quanto alcune licenze (ad esempio, la CC-BY 4.0) contengono un divieto di apporre tali soluzioni, senza eccezioni; altre licenze non contemplano detto profilo (come la IODL); altre ancora contemplano tale divieto, individuando, tuttavia, soluzioni alternative (es. l'OdBL, che prevede il divieto, ma anche la possibilità alternativa di apporre dette misure, a condizione che una copia del database rimanga accessibile senza restrizioni);
- a specifiche distinzioni sulle modalità di utilizzo delle fonti, prevista in alcune licenze (come l'OdBl rispetto ai "database collettivi"), e/o alla diversa gestione, anche sotto il profilo del "copyleft", del licensing del prodotto identificabile come "derivato" (ad esempio, nell'OdBL per il cd. "produced work" - classico esempio, le mappe rispetto al DB geografico - e nella CDLA per i "results from computational use", ovvero i risultati di una analisi algoritmica di diverse fonti per la produzione di un risultato "diverso").

Questi ultimi aspetti sono peculiari ed esulano dalla necessaria semplificazione funzionale alle tabelle sottese; SI RACCOMANDA, pertanto, di fare comunque riferimento, per eventuali approfondimenti, alla serie di risorse utili indicate nel box, svolgendo, ove necessario, specifiche verifiche.

Risorse utili – par. 6.1

- Joinup Licensing Assistant
- Gestione licenze – data.europa.eu
- CDLA permissive compatibility
- CDLA-Permissive-2.0 Compatibility with Other Licenses
- Compatible Licenses, Creative Commons
- Wiki/cc license compatibility
- Compatibility of Creative Commons Licenses
- Italian Open Data License

6.1.2 Buone pratiche: approccio "open by design"

In linea con quanto sopra descritto, SI RACCOMANDA che le PA adottino nella costituzione, generazione ed acquisizione di dataset un **approccio "open by design"** fin dalla progettazione/commissione, come da art. 6, comma 4 del Decreto e sempre nel rispetto della normativa unionale e nazione in materia di protezione dei dati personali come già richiamato nei paragrafi precedenti (cfr. in particolare par. **4.1** e par. **5.1.2**), con particolare riferimento ai principi di cui all'art. 5 del [**GDPR**] e alla protezione dei dati sin dalla progettazione e per impostazione predefinita ex art. 25 del [**GDPR**], secondo i seguenti principi:

1. In caso di nuova costituzione di un dataset, a titolo esemplificativo, SI RACCOMANDA:
   - di inserire clausole contrattuali utili a definire inequivocabilmente la proprietà del dataset in capo alla PA, accompagnate da indicazioni relative all'effettiva e relativa tutela e fruibilità tecnica nel tempo, nella misura possibile (es. indicazioni titolare nei metadati, uso di formati aperti, etc.);
   - nei limiti del possibile, già in fase di analisi, di verificare se il dataset o le componenti del db presentano profili ostativi alla pubblicazione e, in caso positivo, se sussistano soluzioni tecnologiche e/o logico-architetturali utili a rendere pubblicabile almeno parte del dataset o del db, quali:

- uso di layers o altre analoghe modalità che tengano separati i db di origine;
  - percorribilità di richieste di autorizzazioni ad hoc;
  - …
- in detto contesto, di considerare le finalità per le quali i dati sono stati creati e che eventualmente non consentano di renderli automaticamente disponibili in open data;
- nel rispetto di quanto sopra, di dichiarare fin dall'inizio la licenza con cui si intende pubblicare il dataset e dovranno essere fornite indicazioni precise utili a escludere nei limiti del possibile il riutilizzo di fonti terze non compatibili (v. infra) con la licenza di output (di default, appunto, la CC BY 4.0);
- di richiedere, in ogni caso, la tracciatura precisa delle fonti nel caso utilizzate e la predisposizione della documentazione di supporto utile a rispettarne le eventuali condizioni (inclusa l'attribution stessa);
- di prevedere meccanismi utili a verificare l'eventuale bilanciamento di interessi tra la rinuncia ad una eventuale base dati esterna e la possibilità di modificare la licenza di pubblicazione inizialmente ipotizzata, eventuali soluzioni alternative (v. anche infra) nonché la possibilità di scegliere una diversa soluzione di licensing, che dovrà quindi essere motivata secondo i criteri sopra descritti.

Tali condizioni potranno essere applicate tanto nei propri regolamenti interni, così come, appunto, negli accordi negoziali con consulenti e/o fornitori. SI RACCOMANDA, inoltre, di prendere in considerazione tali buone pratiche non solo quando l'oggetto specifico dell'attività dell'ente sia la costituzione del dataset o del DB, ma anche quando il dataset o il DB costituisca un elemento di una attività più ampia di cui è comunque parte essenziale.

In generale, facendo salvo quanto sopra anche con riferimento alle fonti "terze", per i dati che fanno riferimento anche a fonti esterne (per esempio, progetti con altre Pubbliche amministrazioni), dovranno - se possibile già in fase di progettazione - essere verificate le condizioni di riutilizzo di tali fonti; al riguardo, si RACCOMANDA la predisposizione di un report utile a identificare:

- la corretta "titolarità" dei dati (e titolo del relativo riutilizzo);
- le eventuali situazioni di incompatibilità bloccanti una redistribuzione;

- la licenza aperta nel caso adottabile, o le alternative adottabili.
- le alternative tecniche di riutilizzo eventualmente meno "condizionate": a titolo di esempio, si pensi nel contesto dei dati territoriali al ricorso a layers contenenti db del tutto separati rispetto ad un mashup),

procedendo nella costituzione ed evoluzione dei dati secondo step consapevoli, che prendano in considerazione licenze delle fonti esterne e modalità di uso al momento della loro adozione (e non in fase di pubblicazione dei nuovi dati).

Quanto alla fase di pubblicazione, SI RACCOMANDA (eventualmente predisponendo anche in questo caso una apposita check list) di:

- curare la verifica delle condizioni caso per caso richieste dalle licenze "terze" coinvolte;
- rendere la licenza apposta chiaramente individuabile;
- cercare di rendere il rispetto delle condizioni di attribution semplici e di semplice gestione, anche progressiva.

Rispetto a detto ultimo profilo, si rammenta infatti che la clausola di attribution di cui alla CC BY 4.0 prevede:

"*retain the following if it is* **supplied by the Licensor** *with the Licensed Material:*

a) *identification of the creator(s) of the Licensed Material and any others designated to receive attribution, in any reasonable manner requested by the Licensor (including by pseudonym if designated);*
b) *a copyright notice;*
c) *a notice that refers to this Public License;*
d) *a notice that refers to the disclaimer of warranties;*
e) *a URI or hyperlink to the Licensed Material to the extent reasonably practicable*".

La CC-BY 4.0 permette, infatti, di rispettare dette informazioni **nella misura scelta dal licenziante, peraltro in qualsivoglia "forma ragionevole**"; conseguentemente, SI RACCOMANDA di convergere verso una soluzione di attribuzione ove possibile ancora più standardizzata, limitandosi alla mera richiesta di richiamare il nome dell'Ente (come titolare in caso di ripubblicazione di un set di dati non modificato o come fonte di origine in caso di mashup/evoluzione) e precisare se siano state apportate o meno modifiche. A tale proposito, il nome dell'Ente da utilizzare deve corrispondere al nome ufficiale registrato nell'*Indice dei domicili*

digitali della Pubblica Amministrazione e dei Gestori di Pubblici Servizi[50]. Tale nome è anche desumibile dai metadati in quanto nei profili dei metadati indicati nel par. **4.6** sono previsti specifici elementi per indicare il titolare dei dati.

Nei metadati, inoltre, è possibile indicare la licenza applicata ai dati descritti facendo riferimento al relativo vocabolario controllato.

6.2 Tariffazione

Ferma restando la condivisione dei dati tra pubbliche amministrazioni per finalità istituzionali, che avviene esclusivamente a titolo gratuito, salvo per la prestazione di elaborazioni aggiuntive, come previsto dall'art. 50 del **CAD**, il Decreto stabilisce come regola preminente che **i dati siano resi disponibili gratuitamente**; tuttavia, è possibile richiedere per il riutilizzo dei dati i costi marginali sostenuti effettivamente per la riproduzione, la messa a disposizione e la divulgazione dei dati, nonché per l'anonimizzazione di dati personali o per le misure adottate per proteggere le informazioni commerciali a carattere riservato.

REQUISITO 22: dlgs36-2006/opendata/req/conditions/marginal-costs

I dati DEVONO essere resi disponibili per il riutilizzo gratuitamente, salvo eventuale applicazione dei costi marginali effettivamente sostenuti per la riproduzione, la messa a disposizione e la divulgazione dei dati, nonché per l'anonimizzazione di dati personali o per le misure adottate per proteggere le informazioni commerciali a carattere riservato.

Per il calcolo dei costi marginali si può fare riferimento alle indicazioni, tuttora valide, fornite nella Comunicazione della Commissione Europea 2014/C - 240/01 (v. box "Risorse utili" alla fine del paragrafo 6.4) a cui si rimanda per ulteriori approfondimenti. La Comunicazione, in particolare, indica che nell'ambiente in linea, la totalità del corrispettivo potrebbe essere limitata ai costi collegati direttamente alla manutenzione e al funzionamento dell'infrastruttura (ossia della banca dati elettronica) nella misura necessaria a riprodurre il documento e a metterlo a disposizione di un riutilizzatore in più.

Stante quanto sopra, considerate l'esiguità e la tendenza al ribasso dei costi operativi medi di una banca dati, la Comunicazione stima che l'importo possa essere probabilmente **prossimo allo zero**. Da qui la conclusione che **per i documenti digitali divulgati per via elettronica si**

[50] https://indicepa.gov.it

raccomanderebbe una politica di costo zero, quindi la messa a disposizione a titolo gratuito, fermi restando i costi marginali richiesti per una eventuale rielaborazione o anonimizzazione dei dati.

Quanto finora descritto non si applica ad alcuni casi specifici per i quali è possibile determinare tariffe superiori ai costi marginali, ovvero:

1. biblioteche, comprese quelle universitarie, musei e archivi;
2. amministrazioni e organismi di diritto pubblico che devono generare utili per coprire una parte sostanziale dei costi inerenti allo svolgimento dei propri compiti di servizio pubblico;
3. imprese pubbliche.

Nei tre casi appena indicati, può essere applicato quanto indicato nel Requisito 23 che segue.

REQUISITO 23: dlgs36-2006/opendata/req/conditions/reasonable-return
Nel caso in cui sia richiesto il pagamento di un corrispettivo, il totale delle entrate provenienti dalla fornitura e dall'autorizzazione al riutilizzo dei documenti in un esercizio contabile NON PUÒ superare i costi marginali del servizio reso (comprendenti i costi di raccolta, produzione, riproduzione, diffusione, archiviazione dei dati, conservazione e gestione dei diritti e, ove applicabile, di anonimizzazione dei dati personali e delle misure adottate per proteggere le informazioni commerciali a carattere riservato), maggiorati di un utile ragionevole sugli investimenti.

Il significato di "*utile ragionevole sugli investimenti*" viene esplicitato dal Decreto stesso: esso corrisponde ad "*una percentuale della tariffa complessiva, in aggiunta a quella necessaria per recuperare i costi ammissibili (costi marginali), non superiore a cinque punti percentuali oltre il tasso di interesse fisso della BCE*" (cfr. art. 2, comma 1, lettera i-bis)).

Nel caso di biblioteche, musei e archivi, la Direttiva suggerisce, inoltre, che, tenendo conto delle loro peculiarità, nel calcolare l'utile ragionevole sugli investimenti, possano essere presi in considerazione i prezzi praticati dal settore privato per il riutilizzo di documenti identici o simili.

La Comunicazione della Commissione Europea citata innanzi fornisce indicazioni anche per il calcolo delle tariffe in base al metodo del recupero dei costi. Rispetto a quanto indicato nella Comunicazione bisogna tenere presente che la Direttiva ha introdotto il riferimento specifico a

"un esercizio contabile" al posto del generico "periodo contabile adeguato" indicato nelle versioni precedenti e ha introdotto altresì il concetto di "utile ragionevole sugli investimenti", come sopra specificato, invece del "congruo utile sugli investimenti". Tenendo in considerazione queste differenze, il metodo proposto nella Comunicazione al par. 4.2 può essere un utile supporto per il calcolo delle tariffe.

Nei casi indicati ai precedenti punti 2. e 3., vale quanto indicato nel Requisito 24 che segue.

REQUISITO 24: dlgs36-2006/opendata/req/conditions/parameters
L'importo totale delle tariffe DEVE essere calcolato in base a parametri oggettivi, trasparenti e verificabili ed è determinato secondo il criterio del costo marginale del servizio con decreti dei Ministri competenti, di concerto con il Ministro dell'economia e delle finanze sentita l'Agenzia per l'Italia digitale.

Con Decreto del Ministro dell'economia e delle finanze è anche definito e aggiornato periodicamente l'elenco delle pubbliche amministrazioni e degli organismi di diritto pubblico di cui al punto 2, cioè di quegli enti autorizzati a determinare tariffe superiori ai costi marginali in quanto devono generare utili per coprire una parte sostanziale dei costi inerenti allo svolgimento dei propri compiti di servizio pubblico. Tale elenco è pubblicato nei siti istituzionali degli enti interessati.

Il Decreto specifica anche che, rispetto a quanto indicato innanzi, sono fatte salve specifiche disposizioni normative relative, in particolare, alla riutilizzazione commerciale di documenti, dati e informazioni catastali ed ipotecarie, di cui all'art. 1, commi 370, 371 e 372 della legge 30 dicembre 2004, n. 311 e s.m.i., e alla circolazione delle informazioni concernenti gli immobili di cui all'art. 5, comma 4-bis del decreto-legge 13 maggio 2011, n. 70, convertito, con modificazioni, dalla legge 12 luglio 2011, n. 106.

Per quanto riguarda le attività di riutilizzazione commerciale di documenti, dati e informazioni catastali ed ipotecarie, tali attività sono consentite nel rispetto della normativa in materia di protezione dei dati personali.

La norma differenzia tra acquisizione originaria di documenti, dati ed informazioni catastali e acquisizione originaria di documenti, dati ed informazioni ipotecarie.

Nel primo caso, i riutilizzatori commerciali autorizzati devono corrispondere un importo fisso annuale determinato con decreto del Ministro dell'economia e delle finanze, mentre nel

secondo caso devono corrispondere i tributi previsti maggiorati nella misura del 20 per cento. L'importo fisso annuale e la percentuale di aumento possono comunque essere rideterminati annualmente con decreto del Ministro dell'economia e delle finanze anche tenendo conto dei costi complessivi di raccolta, produzione e diffusione di dati e documenti sostenuti dall'Agenzia delle Entrate, maggiorati di un adeguato rendimento degli investimenti e dell'andamento delle relative riscossioni. Con decreto del Ministro dell'economia e delle finanze sono individuate le categorie di ulteriori servizi telematici che possono essere forniti dall'Agenzia delle Entrate esclusivamente ai riutilizzatori commerciali autorizzati a fronte del pagamento di un corrispettivo da determinare con lo stesso decreto.

Ricordiamo che, per agevolare la circolazione delle informazioni concernenti gli immobili, è abolito il divieto di riutilizzazione commerciale dei dati ipotecari e catastali ed è consentito, quindi, il riutilizzo dei documenti, dei dati e delle informazioni catastali e ipotecari a fini commerciali o non commerciali diversi dallo scopo iniziale nell'ambito dei compiti di servizio pubblico per i quali i documenti sono stati prodotti, fermo restando il rispetto della normativa in materia di protezione dei dati personali (art. 5, comma 4-bis, decreto legge 13 maggio 2011, n. 70); inoltre, è comunque consentita la fornitura di documenti, dati e informazioni da parte dell'Agenzia delle Entrate, in formato elaborabile, su base convenzionale, secondo modalità, tempi e costi da stabilire con provvedimento del direttore dell'Agenzia delle Entrate.

REQUISITO 25: dlgs36-2006/opendata/req/conditions/local

Nel caso di enti territoriali ed enti e organismi pubblici diversi da quelli indicati, gli importi delle tariffe, calcolati sulla base dei criteri indicati innanzi, e le relative modalità di versamento sono determinati con disposizioni o atti deliberativi dell'ente titolare, sentita l'Agenzia per l'Italia Digitale.

REQUISITO 26: dlgs36-2006/opendata/req/conditions/publication

Le condizioni applicabili al riutilizzo dei dati e l'effettivo ammontare delle tariffe applicate, compresa la base di calcolo utilizzata per tali tariffe e gli elementi presi in considerazione nel calcolo di tali tariffe, DEVONO essere pubblicati sui siti istituzionali di pubbliche amministrazioni, organismi di diritto pubblico e imprese pubbliche competenti, previa comunicazione ad AgID.

Il riutilizzo DEVE essere **sempre gratuito** nel caso delle **serie di dati di elevato valore** (v. par. **4.3**), con le esclusioni indicate nel **REQUISITO 8**, e nel caso dei **dati della ricerca** (v. par. **4.4**).

6.3 Non discriminazione

Il Decreto stabilisce che le condizioni poste per il riutilizzo (v. par. **6.1**) non debbano comportare discriminazioni per categorie analoghe di riutilizzo, compreso il riutilizzo transfrontaliero. A tale riguardo, il divieto di discriminazioni, per esempio, non deve impedire lo scambio di informazioni tra enti pubblici a titolo gratuito nell'ambito dei loro compiti di servizio pubblico, come peraltro stabilito dall'art. 50 del **CAD**, mentre ai terzi sono applicate tariffe per il riutilizzo degli stessi documenti sulla base delle indicazioni di cui al par. **6.2**.

Se, però, una pubblica amministrazione o un organismo di diritto pubblico riutilizza documenti per attività commerciali che esulano dall'ambito dei propri compiti di servizio pubblico, il Decreto stabilisce che la messa a disposizione dei documenti in questione per tali attività è soggetta alle stesse condizioni e alle medesime tariffe applicate agli altri soggetti che riutilizzano quei documenti.

La Direttiva suggerisce altresì che può essere anche seguita una **politica di tariffe differenziate** per il riutilizzo a fini commerciali e non commerciali.

6.4 Accordi di esclusiva

Si devono evitare, per quanto possibile, accordi di esclusiva tra enti pubblici e partner privati, facendo in modo, sulla base dell'art. 11 del Decreto, che i documenti delle pubbliche amministrazioni, degli organismi di diritto pubblico, delle imprese pubbliche e delle imprese private, possano essere riutilizzati da tutti gli operatori interessati alle condizioni previste dal Decreto stesso, anche qualora uno o più soggetti stiano già procedendo allo sfruttamento di prodotti a valore aggiunto basati su tali documenti. A tale riguardo, il citato art. 11 del Decreto precisa che eventuali contratti o accordi siglati tra enti pubblici e partner privati non danno a questi ultimi diritti esclusivi di utilizzo di documenti detenuti dagli enti pubblici.

In alcuni casi, tuttavia, può essere necessario concedere un diritto esclusivo di riutilizzare determinati documenti del settore pubblico, al fine di garantire un servizio di interesse economico generale.

Gli enti pubblici possono anche emanare specifiche disposizioni che, pur non concedendo espressamente un diritto esclusivo, limitino la disponibilità di riutilizzo di documenti da parte di soggetti diversi da coloro che partecipano agli accordi di cui all'art. 11 del Decreto.

In entrambi i casi, gli accordi o le disposizioni devono essere soggetti ad una valutazione periodica con cadenza almeno triennale per verificare la validità del motivo alla base dell'attribuzione dei diritti esclusivi o l'effetto delle disposizioni.

Nel caso di accordi di cooperazione tra biblioteche, ivi comprese le biblioteche universitarie, musei, archivi e soggetti privati che prevedano la digitalizzazione di risorse culturali garantendo diritti di esclusiva a partner privati, la Direttiva rileva che la prassi ha evidenziato che tali partenariati pubblico-privato possono agevolare un valido utilizzo delle opere culturali e nel contempo accelerare l'accesso dei cittadini al patrimonio culturale.

La Direttiva fa presente, inoltre, che se un diritto esclusivo riguarda la digitalizzazione di risorse culturali, potrebbe essere necessario un certo periodo di esclusiva per dare al partner privato la possibilità di recuperare il suo investimento. Tale periodo dovrebbe tuttavia essere limitato nel tempo ed essere il più breve possibile, al fine di rispettare il principio secondo cui i materiali di dominio pubblico dovrebbero rimanere tali una volta digitalizzati.

Il Decreto ha stabilito che questo periodo **non debba eccedere di norma i sette anni** e che, nel caso di durata superiore, tale durata debba essere soggetta a riesame nel corso dell'ottavo anno e, laddove necessario, successivamente ogni cinque anni.

Nell'ambito degli accordi di cui sopra, sussiste il diritto delle pubbliche amministrazioni e degli organismi di diritto pubblico interessati di ricevere, a titolo gratuito, una copia delle risorse culturali digitalizzate come parte dell'accordo stesso, copia che è resa disponibile per il riutilizzo al termine del periodo di esclusiva.

Gli enti pubblici possono anche emanare specifiche disposizioni che, pur non concedendo espressamente un diritto esclusivo, limitino la disponibilità di riutilizzo di documenti da parte di soggetti diversi da coloro che partecipano all'accordo.

A prescindere dal tipo di accordo (se, cioè, riguardi o meno la digitalizzazione di beni culturali), il Decreto prevede che gli accordi di esclusiva e i relativi termini debbano essere trasparenti e pubblicati nei siti istituzionali, nel caso di accordi e disposizioni che non riguardino la digitalizzazione di beni culturali, almeno due mesi prima che abbiano effetto.

Il Decreto, infine, dispone le scadenze di diritti di esclusiva già esistenti e conclusi e che non possono essere ulteriormente derogati per mancanza di conformità alle condizioni previste dal Decreto stesso. In particolare:

- i diritti di esclusiva esistenti al 17 luglio 2013 e conclusi da pubbliche amministrazioni o da organismi di diritto pubblico cessano alla scadenza del contratto e comunque il 18 luglio 2043, ove la scadenza del contratto sia successiva a tale data;
- i diritti di esclusiva esistenti al 16 luglio 2019 e conclusi da imprese pubbliche cessano alla scadenza del contratto e comunque il 17 luglio 2049, ove la scadenza del contratto sia successiva a tale data.

Risorse utili – parr. 6.2 e 6.4

- Comunicazione della Commissione Europea 2014/C 240/01 - Orientamenti sulle licenze standard raccomandate, i dataset e la tariffazione del riutilizzo dei documenti - **Parte 4**
- *Piano nazionale di digitalizzazione del patrimonio culturale 2022-2023* e relative Linee Guida, Ministero della Cultura
- *Linee Guida per l'acquisizione, la circolazione e il riuso delle riproduzioni dei beni culturali in ambiente digitale*, Ministero della Cultura

Capitolo 7
Pubblicazione e strumenti di ricerca

L'art. 9 del Decreto contiene le previsioni normative relative alle modalità per facilitare la ricerca di dati e documenti resi disponibili per il riutilizzo, con l'indicazione dei due strumenti nazionali a tale scopo deputati, ovvero il portale nazionale dei dati aperti dati.gov.it e il Repertorio Nazionale dei Dati Territoriali.

Quanto alla pubblicazione dei dati, la Direttiva e il Decreto hanno individuato nelle API (Application Programming Interface) lo strumento per rendere disponibili i dati dinamici e le serie di dati di elevato valore (v. art. 6, comma 5 e art. 12-bis del Decreto).

Il presente capitolo, oltre alle indicazioni relative alle disposizioni di cui sopra, contiene anche ulteriori indicazioni non vincolanti - in quanto non derivanti da specifiche disposizioni normative - per la pubblicazione dei dati, già presenti nelle *Linee Guida per la valorizzazione del patrimonio informativo pubblico*.

7.1 Pubblicazione dei dati

7.1.1 API

Per i dati dinamici e le serie di dati di elevato valore, il Decreto prevede che i dati debbano essere resi disponibili attraverso un'interfaccia per programmi applicativi (API), che, come da definizione, è un insieme di funzioni, procedure, definizioni e protocolli per la comunicazione da macchina a macchina e lo scambio ininterrotto di dati. Questo non impedisce, ovviamente, che i titolari dei dati possano continuare a rendere disponibili i dati anche in modalità diverse. Del resto, lo stesso art. 50-ter, comma 7 del **CAD** prevede che le PA possano "*continuare a utilizzare anche i sistemi di interoperabilità già previsti dalla legislazione vigente*".

La Direttiva evidenzia che per la creazione e l'impiego di API sia necessario basarsi su alcuni principi: disponibilità, stabilità, manutenzione per tutto il ciclo di vita, uniformità di utilizzo e delle norme, facilità d'uso e sicurezza.

In tema di API, con la Determinazione di AgID n. 547/2021 sono state adottate le "*Linee Guida sull'interoperabilità tecnica delle Pubbliche Amministrazioni*" [**LG-INT**] e le "*Linee Guida Tecnologie e standard per la sicurezza dell'interoperabilità tramite API dei sistemi informatici*" [**LG-SIC**], che definiscono le modalità con cui le Pubbliche Amministrazioni implementano le proprie API, quale elemento

tecnologico di base del Modello di Interoperabilità (ModI), e individuano le soluzioni tecniche idonee a garantire l'autenticazione dei soggetti coinvolti e la protezione, l'integrità e la riservatezza dei dati scambiati nelle interazioni tra sistemi informatici della pubblica amministrazione e di questi con i sistemi informatici di soggetti privati per il tramite di API.

Le API che il Decreto indica di utilizzare per rendere disponibili dati dinamici e serie di dati di elevato valore DEVONO essere sviluppate e pubblicate in conformità con le Linee Guida citate innanzi.

REQUISITO 27: dlgs36-2006/opendata/req/publication/api
Le API sviluppate per rendere disponibili i dati per il riutilizzo DEVONO essere conformi alle "Linee Guida sull'interoperabilità tecnica delle Pubbliche Amministrazioni" e le "Linee Guida Tecnologie e standard per la sicurezza dell'interoperabilità tramite API dei sistemi informatici", adottate con la Determinazione di AgID n. 547/2021.

Dette Linee Guida individuano il Catalogo delle API quale componente, unica e centralizzata, che assicuri alle parti coinvolte nel rapporto di erogazione e fruizione la consapevolezza sulle API disponibili, e per esse, i livelli di servizio dichiarati. Tale Catalogo, in applicazione dell'art. 50 ter del **CAD**, è reso disponibile attraverso la Piattaforma Digitale Nazionale Dati (PDND) a cui si applicano le indicazioni delle "*Linee Guida sull'infrastruttura tecnologica della Piattaforma Digitale Nazionale Dati per l'interoperabilità dei sistemi informativi e delle basi di dati*" adottate con la Determinazione di AgID n. 627/2021 [**LG-PDND**].

La coerenza con le Linee Guida suddette garantisce anche l'adempimento di quanto previsto nel Regolamento (UE) [**REG-HVD**] relativo ai dati di elevato valore (v. par. **4.3**), in particolare con la previsione di cui all'art. 3(2) circa le condizioni di utilizzo delle API e i criteri di qualità del servizio.

Il paragrafo del Regolamento citato dispone, infatti, che debbano essere stabiliti e pubblicati i termini di utilizzo delle API e i criteri di qualità dei servizi in relazione a prestazione, capacità e disponibilità. Tale disposizione può essere estesa anche oltre l'ambito di applicazione del suddetto Regolamento, applicandola in modo generalizzato in modo che i titolari dei dati assumano impegni precisi e non ci si affidi al criterio del *best-effort*. Nel caso dei servizi di rete INSPIRE, i criteri di qualità sono definiti con il Regolamento (CE) n. 976/2009 e s.m.i. In generale, occorre fare riferimento ai requisiti e alle indicazioni sulla qualità dei servizi definiti nelle Linee Guida sull'interoperabilità tecnica delle Pubbliche Amministrazioni.

Considerato che, come evidenziato dalla Direttiva, è, in generale, vantaggioso riutilizzare e condividere i dati tramite un impiego adeguato di API, perché possono aiutare gli sviluppatori e le start-up a creare nuovi servizi e prodotti e rappresentano, inoltre, un elemento fondamentale della strutturazione di ecosistemi di valore attorno a un patrimonio di dati spesso inutilizzato, è fortemente raccomandato l'utilizzo di API per rendere disponibili per il riutilizzo anche quelle tipologie di dati per le quali il Decreto non prevede tale disposizione.

In caso di serie storiche di dati, la pubblicazione dovrebbe riguardare non solo i dati più recenti o attuali, ma tutta la successione in modo da non disperdere un patrimonio informativo importante utile, per esempio, per analisi sull'evoluzione dei vari fenomeni di interesse.

Raccomandazione 15: dlgs36-2006/opendata/rec/publication/api
Ove possibile, SI RACCOMANDA di utilizzare API conformi al **REQUISITO 27** per rendere disponibili per il riutilizzo tutte le tipologie di dati, non solo quelli dinamici e/o di elevato valore.

Nel caso di dati territoriali, viste le regole tecniche di cui alle Linee Guida INSPIRE per l'implementazione dei servizi di rete di cui all'art. 7, comma 1 del D. Lgs.n. 32/2010 [**D-LGS-32-2010**], **detti servizi di rete possono essere considerati come API** in considerazione del fatto che nell'ambito delle Linee Guida sull'interoperabilità tecnica delle PA, accettando la nomenclatura in uso a livello europeo e più in generale nel contesto internazionale, si utilizza il termine generico API per indicare indifferentemente le Web API, i web service e le API REST.

REQUISITO 28: dlgs36-2006/opendata/req/publication/inspire-ns
Nel caso di dati territoriali, il **REQUISITO 27** è attuato attraverso l'implementazione dei servizi di rete di cui all'art. 11 della Direttiva 2007/2/CE, del Regolamento (CE) n. 976/2009 e delle relative Linee Guida tecniche.

Nel caso dell'applicazione del **REQUISITO 28**, fermo restando, comunque, quanto previsto dal già citato art. 50-ter, comma 7 del **CAD**, detti servizi potrebbero essere documentati anche nel catalogo API disponibile nella PDND.

7.1.1.1 Sicurezza e disponibilità dei dati

L'art. 51, comma 1 del **CAD** demanda a specifiche Linee guida l'individuazione delle soluzioni tecniche idonee a **garantire la protezione, la disponibilità, l'accessibilità, l'integrità e la riservatezza dei dati** e la continuità operativa dei sistemi e delle infrastrutture.

Tali Linee Guida sono state definite nell'ambito del modello di interoperabilità delle pubbliche amministrazioni e corrispondono alle già citate Linee Guida "*Tecnologie e standard per la sicurezza dell'interoperabilità tramite API dei sistemi informatici*" [**LG-SIC**], che definiscono le tecnologie e le loro modalità di utilizzo al fine di garantire la sicurezza delle transazioni digitali realizzate tra e verso le pubbliche amministrazioni che utilizzano le API tramite rete di collegamento informatica.

L'implementazione di API coerenti con il **REQUISITO 27**, pertanto, assicura il rispetto degli adempimenti e la conformità al citato art. 51 del **CAD**.

Nell'ambito della Strategia Cloud Italia (v. box "Risorse utili"), a cui si rimanda, inoltre, al fine di regolamentare l'offerta di servizi cloud disponibili sul mercato per mitigare i rischi sistemici di sicurezza e affidabilità, è stata definita una classificazione dei dati e dei servizi allo scopo di definire un processo di classificazione dei dati, in base al danno che una loro compromissione potrebbe provocare al sistema Paese (strategici, critici e ordinari). Il risultato della classificazione, si legge nella strategia, consente di uniformare e guidare il processo di migrazione al Cloud della PA. Le classi individuate sono:

- **Strategico**: dati e servizi la cui compromissione può avere un impatto sulla sicurezza nazionale;
- **Critico**: dati e servizi la cui compromissione potrebbe determinare un pregiudizio al mantenimento di funzioni rilevanti per la società, la salute, la sicurezza e il benessere economico e sociale del Paese;
- **Ordinario**: dati e servizi la cui compromissione non provochi l'interruzione di servizi dello Stato o, comunque, un pregiudizio per il benessere economico e sociale del Paese.

Risorse utili – par. 7.1.1

- OpenAPI Checker, Developers Italia
- Guide tecniche INSPIRE per i servizi di rete
- *Evaluation of Application Programming Interfaces for INSPIRE*, Publications Office of the European Union, 2022
- OGC API – Features as an INSPIRE download service
- OGC SensorThings API as an INSPIRE download service
- Strategia Cloud Italia

7.1.2 Elementi architetturali

I principali livelli architetturali che compongono una soluzione per la pubblicazione e l’interrogazione di dati aperti possono essere istanziati in diverso modo a seconda delle capacità economiche e tecniche delle amministrazioni, nonché della qualità del servizio che si vuole offrire agli utenti. Si distinguono due livelli: livello di front-end e livello infrastrutturale. Il livello di front-end consiste di una parte di presentazione che può essere sia un sito Web, sia una sezione in un sito esistente. In questa parte rientrano tutti quegli strumenti che consentono di dare massima visibilità ai dati disponibili e di interagire in maniera “user-friendly” con gli utenti stessi, per esempio per capire quali dati siano di loro interesse, quali nuovi dati siano richiesti, quali suggerimenti vogliano dare per migliorare anche la qualità dei dati. Il livello di presentazione si completa con l’interfaccia di accesso via web per interrogazioni puntuali sui dati e sui metadati. L’interfaccia ha come obiettivo quello di aumentare l’interazione machine-to-machine attraverso il dispiegamento di una piattaforma di esposizione dati basata su API di servizio (o Open Data Service). Nel caso di dati dei livelli 4 e 5 del modello per i dati di cui all’Allegato A, l’interfaccia di accesso via web è rappresentata dallo SPARQL endpoint.

In generale, SI RACCOMANDA di:

- assegnare ai dataset nomi autoesplicativi per comprenderne il principale contenuto;
- fornire, ove possibile, descrizioni testuali dei dataset;
- mettere in evidenza la licenza in uso in forma “human and machine-readable”;
- fornire, ove possibile, strumenti di visualizzazione e navigazione, anche georiferita, dei dati, che possano facilitare la lettura degli stessi;

- fornire, ove possibile, statistiche di uso, accesso e produzione;
- fornire notifiche di cambiamenti nel sito web, di aggiornamenti ai dataset (per es., RSS feed);
- fornire strumenti per rendere le interrogazioni più agevoli, anche per utenti non del tutto esperti. Nel caso dei dati dei livelli 4 e 5 non si può pubblicare solo dataset RDF ma è bene mettere in evidenza la presenza dello SPARQL endpoint (cioè, un servizio Web che accetta interrogazioni SPARQL, le risolve e restituisce i risultati in output), pubblicando il link di accesso, fornendo altresì un ampio insieme di "query" di esempio che possono essere eseguite producendo risultati disponibili in diversi formati di più comune utilizzo soprattutto nell'ambito delle comunità open data (per es., CSV, JSON, XML).

Nei casi di amministrazioni di minori dimensioni o amministrazioni che non siano nelle condizioni di poter fornire un servizio con le caratteristiche sopra elencate, si consiglia di implementare azioni di sussidiarietà verticale (ad esempio, i comuni di medio-piccole dimensioni possono riferirsi alla Regione di appartenenza) o di unirsi in iniziative comuni (v. paragrafo **5.1.1.1**).

Il livello infrastrutturale è rappresentato dall'infrastruttura che ospita i dati e i metadati. Nel caso di dati aperti, tenuto conto della loro natura intrinseca, ovvero di dati tipicamente non riferibili a singole persone e per i quali solitamente non si richiede il soddisfacimento di specifici requisiti di protezione dei dati personali, tecnologie basate sul paradigma del cloud computing pubblico possono essere facilmente impiegabili al fine di ospitare le infrastrutture per la pubblicazione di dati aperti.

In tema di cloud, la Strategia Cloud Italia (v. box "Risorse utili"), già citata innanzi, definita dal Dipartimento per la trasformazione digitale della Presidenza del Consiglio dei Ministri e dall'Agenzia per la cybersicurezza nazionale (ACN), contiene gli indirizzi strategici per il percorso di migrazione verso il cloud di dati e servizi digitali della Pubblica Amministrazione. Per l'attuazione della strategia, AgID ha adottato, con la Determinazione n. 628/2021, il Regolamento che definisce i requisiti minimi per le infrastrutture digitali e le caratteristiche e le modalità di qualificazione e migrazione dei servizi cloud.

Risorse utili – par. 7.1.2

- Strategia Cloud Italia
- *Regolamento recante i livelli minimi di sicurezza, capacità elaborativa, risparmio energetico e affidabilità delle infrastrutture digitali per la PA e le caratteristiche di qualità, sicurezza, performance e scalabilità, portabilità dei servizi cloud per per la pubblica amministrazione, le modalità di migrazione, nonché le modalità di qualificazione dei servizi cloud per la pubblica amministrazione*, adottato con la Determinazione AgID n. 628/2021 del 15 dicembre 2021.

7.1.3 Identificatori univoci e persistenti

Nei requisiti per i dati della ricerca volti a rendere tali dati conformi ai principi FAIR si è fatto riferimento più volte a identificatori univoci e persistenti. Tali identificatori (che è buona prassi applicare a tutti i dati e non solo a quelli della ricerca) sono generalmente rappresentati dagli URI (Uniform Resource Identifier), una sequenza di caratteri che identifica una risorsa astratta o fisica. Essi sono utilizzati nei linked data (v. par. **5.1.4**) per risolvere il problema dell'identità.

Gli URI DEVONO essere persistenti e dereferenziabili. Una politica per garantire URI persistenti e fornire aspetti di naming è proposta in uno studio dalla Commissione Europea (v. box "Risorse utili" in calce). Facendo riferimento a tale documento, per la creazione di URI persistenti sono da evitare quelli che contengano:

- nome del progetto/ufficio/unità amministrativa che detiene la risorsa per evitare problemi derivanti dalla fine del progetto stesso o fusioni o chiusure di uffici nell'organizzazione;
- numeri di versione;
- identificatori esistenti che in passato siano stati utilizzati per identificare risorse differenti;
- riferimenti generati in modo automatico e incrementale a meno che non vi sia la garanzia che il processo non venga mai più ripetuto o, se ripetuto, generi sicuramente gli stessi identificatori per gli stessi dati di input;
- stringhe rappresentanti "query" a database;
- estensione del file.

Sono, invece, da ritenersi buone pratiche le seguenti:

- strutturare l'URI come segue:

  `http://{dominio}/{tipo}/{concetto}/{riferimento}`

- includere nell' URI i seguenti elementi:
  - *dominio*: il dominio Web su cui reperire la risorsa
  - *tipo*: l'elemento che specifica il tipo di risorsa. Dovrebbe poter assumere un numero limitato di valori come "doc" se la risorsa identificata è un documento descrittivo, "set" se la risorsa è un dataset, "id" o "item" se la risorsa è un oggetto del mondo reale
  - *concetto*: il tipo di un oggetto del mondo reale
  - *riferimento*: lo specifico elemento, termine o concetto che rappresenta la risorsa
- costruire URI per più formati al fine di identificare al meglio la risorsa
- collegare tra loro le rappresentazioni multiple della stessa risorsa
- implementare il codice di risposta 303 per gli oggetti del mondo reale e risorse in generale (si veda di seguito "content negotiation" e "dereferenziazione" degli URI)
- utilizzare servizi dedicati.

Nella gestione degli URI è opportuno utilizzare il meccanismo cosiddetto di "*content negotiation*" che consente di rendere disponibile, allo stesso URI, diverse rappresentazioni di una risorsa in caso di molteplici rappresentazioni possibili (per es. URI che rappresentano sia entità del web semantico sia risorse web). Analogamente, è una buona prassi utilizzare sempre URI "deferenziabili", come già indicato prima, cioè URI che restituiscono una rappresentazione web (ad es. una pagina web informativa) di una risorsa.

Risorse utili – par. 7.1.3

- Cool URIs for the Semantic Web, W3C
- Study on persistent URIs, with identification of best practices and recommendations on the topic for the MSs and the EC, programma ISA, Commissione Europea

7.2 Strumenti per la ricerca

Il Decreto individua come strumenti per la ricerca dei dati il **catalogo nazionale dei dati aperti**[51] e, per i dati territoriali, il **Repertorio Nazionale dei Dati Territoriali (RNDT)**[52] di cui all'art. 59 del **CAD**, entrambi gestiti da AgID.

Ai sensi dell'articolo 9 del Decreto, **il portale nazionale dei dati aperti (dati.gov.it) è l'unico riferimento per la documentazione e la ricerca di tutti i dati aperti della pubblica amministrazione**. Esso, in quanto punto di accesso nazionale per i metadati dei dati aperti, **è l'unico ad abilitare il colloquio con l'analogo** portale ufficiale dei dati europei data.europa.eu. Il portale nazionale include i metadati, conformi al profilo DCAT-AP_IT, che descrivono i dati aperti delle amministrazioni secondo quanto indicato al par. **4.6**.

REQUISITO 29: dlgs36-2006/opendata/req/publication/od-portal
Le amministrazioni sono tenute a inserire e a mantenere aggiornati nel portale dati.gov.it, attraverso le modalità di alimentazione previste dal catalogo, i metadati dei dati, ad esclusione di quelli territoriali.

I dati, il cui riferimento è pubblicato sul portale nazionale, rimangono presso il titolare del dato che conserva la responsabilità della loro correttezza, tenuta, gestione, aggiornamento e divulgazione a livello nazionale.

REQUISITO 30: dlgs36-2006/opendata/req/publication/rndt
I dati territoriali DEVONO essere documentati esclusivamente presso il Repertorio Nazionale dei Dati Territoriali (RNDT) che, in maniera automatizzata, si occupa dell'allineamento con il portale nazionale dei dati aperti dati.gov.it.

L'integrazione dei due cataloghi avviene attraverso la specifica GeoDCAT-AP che fornisce una sintassi RDF per i metadati INSPIRE e ISO 19115 (profilo core) e la sua estensione italiana che è stata definita con la *Guida nazionale per l'implementazione della specifica GeoDCAT-AP* (v. box "Risorse utili").

Nella citata Guida sono definite, tra l'altro, le due regole principali così formulate:

[51] https://dati.gov.it
[52] https://geodati.gov.it

1. I dati territoriali, anche quando siano resi disponibili secondo il paradigma open data, **DEVONO essere documentati ESCLUSIVAMENTE nel RNDT** secondo le regole nazionali sui metadati dei dati territoriali e le relative guide operative;
2. Il RNDT garantirà l'accesso ai dati territoriali "di tipo aperto" anche nel catalogo nazionale dei dati aperti (dati.gov.it), secondo lo standard DCAT-AP_IT, attraverso GeoDCAT-AP e sulla base delle corrispondenze definite nel documento.

Sulla base di queste regole e delle corrispondenze definite tra i metadati dei due standard di riferimento, i dati territoriali aperti documentati nel RNDT sono resi accessibili attraverso le funzionalità tipiche del catalogo nazionale dei dati aperti, dati.gov.it, senza nessun ulteriore adempimento da parte dei titolari di dati.

Come indicato nelle Linee Guida RNDT [**LG-RNDT**], il Repertorio, in quanto punto di accesso nazionale per i metadati dei dati territoriali, provvede a rendere disponibili i metadati anche al geoportale INSPIRE[53] secondo le modalità individuate per l'applicazione di **INSPIRE-DIR**, anche ai fini delle operazioni di monitoraggio di cui alla Decisione di esecuzione (UE) 2019/1372 della Commissione Europea[54].

Nella figura che segue è rappresentato lo schema di coordinamento e integrazione tra i due cataloghi nazionali e la loro interazione con i portali europei.
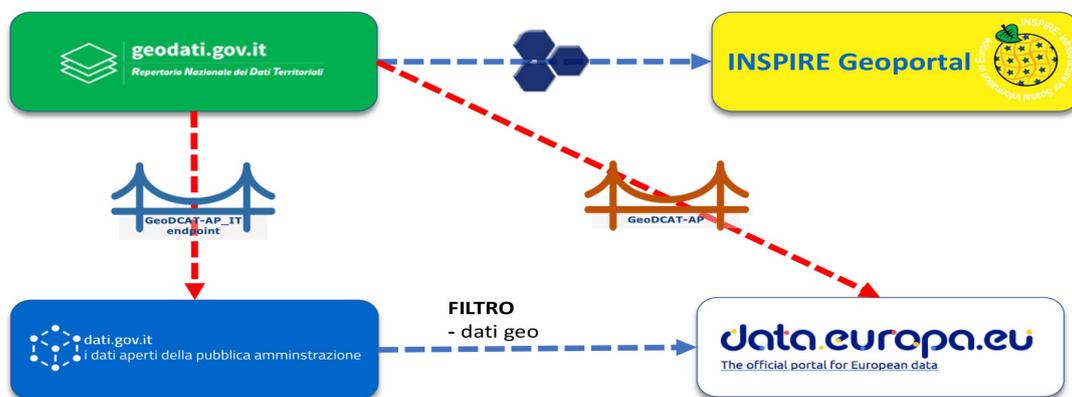
Figura 4 - Interoperabilità tra cataloghi nazionali ed europei

Come indicato nella guida nazionale su GeoDCAT-AP, quanto sopra è riferito alla integrazione dei cataloghi nazionali RNDT e dati.gov.it e all'allineamento delle informazioni in essi

[53] https://inspire-geoportal.ec.europa.eu/

[54] Decisione di esecuzione (UE) 2019/1372 della Commissione del 19 agosto 2019 recante attuazione della direttiva 2007/2/CE del Parlamento europeo e del Consiglio per quanto riguarda il monitoraggio e la comunicazione

contenute; ferma restando la possibilità di un diverso orientamento, SI RACCOMANDA alle amministrazioni di adottare l'approccio nazionale anche nella gestione dei cataloghi "locali".

Analogamente ai dati territoriali, anche per i dati statistici è stata definita un'estensione della specifica DCAT-AP, denominata StatDCAT-AP[55] (v. box "Risorse utili"), allo scopo di garantire la disponibilità dei dati statistici, utilizzando gli standard specifici, come SDMX, anche nei portali generalisti dei dati aperti.

Raccomandazione 16: dlgs36-2006/opendata/rec/publication/statdcat-ap
Al fine di poter documentare i dati statistici, compresi quelli identificati come dati ad elevato valore, nel portale nazionale dati.gov.it secondo il relativo profilo di metadati basato sulla specifica europea DCAT-AP, SI RACCOMANDA di fare riferimento all'estensione StatDCAT-AP.

Risorse utili – par. 7.2

- Guida nazionale per l'implementazione della specifica GeoDCAT-AP
- StatDCAT Application Profile for data portals in Europe

7.2.1 Elenchi delle categorie e modalità di ricerca

Il Decreto dispone che i soggetti destinatari delle presenti Linee Guida debbano individuare le modalità per facilitare la ricerca, anche interlinguistica, dei documenti disponibili per il riutilizzo, insieme ai rispettivi metadati, ove possibile accessibili on-line e in formati leggibili meccanicamente. Tale disposizione è attuata attraverso il rispetto del **REQUISITO 20** e **REQUISITO 21** definiti innanzi.

Il Decreto prevede, inoltre, quanto indicato nel seguente Requisito 31.

REQUISITO 31: dlgs36-2006/opendata/req/publication/categories
I destinatari delle presenti Linee Guida DEVONO pubblicare e aggiornare annualmente nei propri siti istituzionali gli elenchi delle categorie di dati detenuti ai fini del riutilizzo attraverso collegamenti ipertestuali al portale nazionale dati.gov.it.

[55] Attualmente l'estensione StatDCAT-AP è allineata alla versione 1.1 di DCAT-AP

Per adempiere a tale disposizione, considerato che in base ai Requisiti 20 e 21 i dati DEVONO essere documentati nel portale nazionale dei dati aperti dati.gov.it, si può pubblicare nel proprio sito istituzionale, eventualmente in una sezione dedicata agli open data, anche al fine di evitare eventuali duplicazioni, un **collegamento ipertestuale** (anche sfruttando le API disponibili nel portale nazionale), per ciascuna categoria tematica (facendo riferimento ai temi DCAT-AP), ai propri dataset pubblicati nel portale nazionale. Un esempio di URL da inserire nel proprio sito istituzionale è il seguente:

https://dati.gov.it/view-dataset?holder_name=%22Provincia%20Autonoma%20di%20Trento%22&groups=ambiente

dove l'amministrazione titolare è la Provincia Autonoma di Trento e la categoria è "ambiente".

Le categorie tematiche da considerare sono quelle corrispondenti ai temi definiti nell'ambito della specifica DCAT-AP[56] che si riportano di seguito:

- Agricoltura, pesca, silvicoltura e prodotti alimentari;
- Ambiente;
- Economia e finanze;
- Energia;
- Giustizia, sistema giuridico e sicurezza pubblica;
- Governo e settore pubblico;
- Istruzione, cultura e sport;
- Popolazione e società;
- Regioni e città;
- Salute;
- Scienza e tecnologia;
- Tematiche internazionali;
- Trasporti.

Tali temi, in quanto definiti nel profilo di metadati DCAT-AP_IT, sono già utilizzati come classificazione dei dataset nei portali dei dati aperti, a partire da quello nazionale dati.gov.it.

[56] v. il relativo vocabolario controllato disponibile al link https://op.europa.eu/en/web/eu-vocabularies/concept-scheme/-/resource?uri=http://publications.europa.eu/resource/authority/data-theme

Nel caso di dati territoriali aperti, come già indicato, questi vanno documentati esclusivamente nel RNDT che li renderà disponibili anche in dati.gov.it, consentendo, quindi, la creazione del collegamento ipertestuale di cui sopra.

Per la pubblicazione dei metadati nel catalogo nazionale, ci si può avvalere dei **principi di sussidiarietà verticale**, già in precedenza menzionati, o **adottare autonomamente una delle possibili soluzioni** descritte di seguito.

Per quanto riguarda la sussidiarietà verticale, nell'ambito locale, **le Regioni possono assumere il ruolo di aggregatori territoriali**. In sostanza, la Regione, ove possibile, si coordina con le varie amministrazioni che operano nell'ambito territoriale della Regione stessa, raccogliendo le informazioni sui dataset disponibili come dati aperti e assicurando una adeguata frequenza di aggiornamento. Le amministrazioni locali delegano così la Regione all'esposizione dei propri metadati e **possono evitare di richiedere direttamente la raccolta degli stessi da parte del portale nazionale**; quest'ultimo si interfaccia quindi con i soli cataloghi regionali.

Lo stesso modello può applicarsi nei casi di amministrazioni centrali che svolgano un ruolo di "coordinamento" nei riguardi di altre amministrazioni. In questo caso, si richiede che le amministrazioni comunichino tale situazione al portale nazionale durante la fase di richiesta di raccolta.

Per quanto riguarda le possibili soluzioni "autonome", che possono essere adottate anche qualora il modello di sussidiarietà di cui sopra non potesse essere applicato (per es., per mancanza di un catalogo regionale o difficoltà, anche tecniche, di colloquio tra i diversi livelli amministrativi locali), di seguito si riportano alcune di queste possibili soluzioni per la creazione di piattaforme di pubblicazione dei dati.

Soluzione nativa. Viene creato un portale ad-hoc o creata un'apposita sezione di un portale esistente. In questo caso, la creazione non differisce dalla creazione di un sito web classico.

Estensione soluzione CMS esistente. Molto spesso l'amministrazione gestisce già un sito web, realizzato mediante l'uso di un CMS, che vuole estendere con una sezione dedicata agli Open Data. La criticità in questo caso è data dall'aggiunta di una componente semantica all'interno della configurazione del CMS stesso.

Utilizzo di piattaforme esterne. Viene utilizzata una piattaforma che include funzionalità per la catalogazione, visualizzazione, ricerca e interrogazione dei dati. In alcuni casi queste piattaforme sono disponibili in modalità «*as a service*».

Raccomandazione17: dlgs36-2006/opendata/rec/publication/portals
SI RACCOMANDA di non creare tanti portali diversi per singole iniziative ma, ove possibile, di raccordarle per facilitare il reperimento e il riutilizzo dei dati da parte degli utenti finali.

Allegato A
Modello per i dati aperti

Il modello da considerare per produrre e pubblicare dati aperti è quello noto come modello a cinque stelle rappresentato in **Figura 5**, che fornisce una scala per misurare l'apertura dei dati.
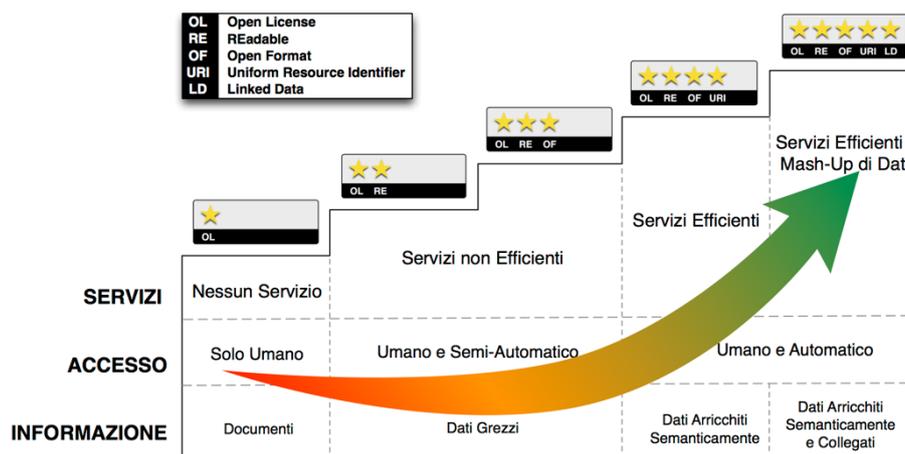
Figura 5 - Modello a cinque stelle per i dati aperti[57]

1. I livelli del modello per i dati aperti

Per ciascun livello, di seguito vengono indicate le caratteristiche principali in termini di informazione, accesso e servizi.

La **Tabella 1** riportata al par. **4.1**, indica, per ciascun formato, il numero di stelle raggiungibile.

1.1 Livello 1 (1 stella)

- **Informazione**: Dati disponibili tramite una licenza aperta e inclusi in documenti leggibili e interpretabili solo grazie a un significativo intervento umano (per es., PDF);
- **Accesso**: Prevalentemente umano, necessario anche per dare un senso ai dati inclusi nei documenti;
- **Servizi**: Solo rilevanti interventi umani di estrazione ed elaborazione dei possibili dati consentono di sviluppare servizi con l'informazione disponibile in questo livello;

[57] Rivisitazione della figura disponibile sul web

1.2 Livello 2 (2 stelle)

- **Informazione:** Dati disponibili in forma strutturata e con licenza aperta. Tuttavia, i formati sono proprietari (per es., Excel) e un intervento umano è fortemente necessario per un'elaborazione dei dati;
- **Accesso:** I programmi possono elaborare i dati ma non sono in grado di interpretarli; pertanto è necessario un intervento umano al fine di scrivere programmi ad-hoc per il loro utilizzo;
- **Servizi:** Servizi ad-hoc che devono incorporare i dati per consentire un accesso diretto via Web agli stessi;

1.3 Livello 3 (3 stelle)

- **Informazione**: Dati con caratteristiche del livello precedente ma in un formato non proprietario (per es., CSV, JSON, geoJSON). I dati sono leggibili da un programma ma l'intervento umano è necessario per una qualche elaborazione degli stessi;
- **Accesso**: I programmi possono elaborare i dati ma non sono in grado di interpretarli; pertanto è necessario un intervento umano al fine di scrivere programmi ad-hoc per il loro utilizzo;
- **Servizi**: Servizi ad-hoc che devono incorporare i dati per consentire un accesso diretto via Web agli stessi;

1.4 Livello 4 (4 stelle)

- **Informazione**: Dati con caratteristiche del livello precedente ma esposti usando standard W3C quali RDF e SPARQL. I dati sono descritti semanticamente tramite metadati e ontologie;
- **Accesso**: I programmi sono in grado di conoscere l'ontologia di riferimento e pertanto di elaborare i dati quasi senza ulteriori interventi umani;
- **Servizi**: Servizi che sfruttano accessi diretti a Web per reperire i dati di interesse;

1.5 Livello 5 (5 stelle)

- **Informazione**: Dati con caratteristiche del livello precedente ma collegati a quelli esposti da altre fonti (per es., Linked Open Data). I dati sono descritti semanticamente tramite metadati e ontologie. Essi seguono il paradigma RDF, in cui alle entità è assegnato un URI univoco sul Web. Nel caso dei Linked Open Data l'intervento umano è minimo o nullo;
- **Accesso**: I programmi sono in grado di conoscere l'ontologia di riferimento e pertanto di elaborare i dati quasi senza ulteriori interventi umani;
- **Servizi**: Servizi che sfruttano sia accessi diretti a Web sia l'informazione ulteriore catturata attraverso i link dei dati di interesse, facilitando il mashup di dati.

Allegato B
Standard di riferimento e formati aperti

1. Standard di riferimento

Come già indicato per i principi FAIR (v. par. **4.4**) e, più in generale, nelle indicazioni fornite nei paragrafi **5.1.4** e **5.1.5** relativi alle fasi, rispettivamente, di arricchimento e di modellazione, per assicurare l'interoperabilità e consentire che dati e metadati possano essere combinati con altri dati e/o strumenti, è necessario, tra l'altro, che vengano utilizzati standard pertinenti, oltre a vocabolari controllati, thesauri e ontologie, riconosciuti auspicabilmente a livello internazionale.

Nel pubblicare dati aperti, quindi, sarebbe opportuno, ove possibile, seguire standard definiti dagli organismi di standardizzazione internazionali, come ISO, W3C, OGC, IETF, o nell'ambito delle attività istituzionali della Commissione Europea. Nel caso in cui non siano disponibili standard a livello internazionale e/o europeo, allora si può fare riferimento a standard e regole tecniche nazionali, anche definiti dalle amministrazioni competenti in funzione dello specifico dominio. Si richiama qui quanto indicato per i modelli dati al par. **5.1.5**.

La tabella che segue riporta l'elenco, non esaustivo, dei principali standard di riferimento. In aggiunta a quelli riportati, sono da considerare i documenti tecnici (come, per es., le Linee Guida) indicati nel **Capitolo 2**.

Risorse utili

- European Catalogue on the ICT standards, Commissione Europea
- Architectures and Standards for Spatial Data Infrastructures and Digital Government, JRC Technical Report, 2020

Acronimo/ abbreviazione	Titolo	Organismo di standardizzazione	Dominio	URL	Note
Rappresentazione delle informazioni					
DataCube	The RDF Data Cube Vocabulary	W3C	Dati statistici	https://www.w3.org/TR/vocab-data-cube/	
EU vocabularies	Risorse (vocabolari controllati, modelli, schemi e ontologie) rese disponibili dall'Ufficio delle Pubblicazioni dell'Unione Europea	Commissione Europea	Vari domini	https://op.europa.eu/en/web/eu-vocabularies/home	
INSPIRE	Infrastructure for Spatial Information in the European Community	Commissione Europea	Dati territoriali	http://inspire.ec.europa.eu/	
ISO 639	Language codes	ISO	Lingua	http://www.iso.org/iso/home/standards/language_codes.htm	
ISO 8601	Date and time format	ISO	Tempo	http://www.iso.org/iso/home/standards/iso8601.htm	
ISO 19100	Serie 19100 "Geographic Information"	ISO	Dati territoriali	https://www.iso.org/committee/54904/x/catalogue/	
NDC	Catalogo nazionale della semantica dei dati	Governo italiano / ISTAT	Vari domini	https://www.schema.gov.it/	
NUTS	Nomenclature of territorial units for statistics	Commissione Europea	Dati territoriali	https://ec.europa.eu/eurostat/web/nuts/background	
SDMX	Statistical Data and Metadata eXchange	SDMX community	Dati statistici	https://sdmx.org/?page_id=5008	SDMX è anche uno Standard ISO (ISO 17369:2013 https://www.iso.org/standard/52500.html

	Regole tecniche dati territoriali	Governo italiano / AgID	Dati territoriali	https://geodati.gov.it/geoportale/datiterritoriali/regole-tecniche	
Metadati					
DCAT	Data Catalog Vocabulary	W3C	Vari domini	http://www.w3.org/TR/vocab-dcat/	
DCAT-AP_IT	DCAT Application Profile (Italia)	AgID	Vari domini	https://docs.italia.it/italia/daf/linee-guida-cataloghi-dati-dcat-ap-it/it/stabile/dcat-ap_it.html	
DCMI	Dublin Core Metadata Initiative	Dublin Core	Vari domini	http://dublincore.org/documents/dcmi-terms/	DCMI è anche disponibile come Standard ISO: - ISO 15836-1:2017 *Information and documentation — The Dublin Core metadata element set — Part 1: Core elements* (https://www.iso.org/standard/71339.html); - ISO 15836-2:2019 *Information and documentation — The Dublin Core metadata element set — Part 2: DCMI Properties and classes* (https://www.iso.org/standard/71341.html)
RNDT	Profilo italiano di metadati per i dati territoriali e relativi servizi	AgID	Dati territoriali	https://agid.github.io/geodocs/rndt-lg/2.0.1/	
Web semantico					
OWL	Ontology Web Language	W3C		https://www.w3.org/TR/owl-features/	
RDF	Resource Description Framework	W3C		https://www.w3.org/TR/rdf11-primer/	
RDFS	RDF Schema	W3C		https://www.w3.org/TR/rdf-schema/	
SPARQL	SPARQL Protocol for RDF	W3C		https://www.w3.org/TR/rdf-sparql-protocol/	

Tabella 7 - Principali standard di riferimento

2. Formati aperti per dati e documenti

Nei sottoparagrafi che seguono vengono elencati i più comuni formati per dati e documenti; un elenco comunque non esaustivo. Per alcuni formati, vengono indicate anche alcune puntuali raccomandazioni tratte dal documento "data.europa.eu – Data quality guidelines", indicato nel box "Risorse utili", a cui si rimanda per ulteriori approfondimenti.

Una raccomandazione generale valida per tutti i formati è la seguente:

- **Utilizza codifiche dei caratteri standard**
  Per garantire che i caratteri siano visualizzati correttamente e avere la massima compatibilità possibile con le applicazioni che elaborano i dati, è sempre necessario utilizzare una codifica standardizzata dei caratteri. In genere, UTF-8 è la codifica utilizzata nel web. È utile, in ogni caso, indicare qual è la codifica dei caratteri utilizzata.

2.1 Formati aperti per i dati

2.1.1 CSV (Comma Separated Values)

È un formato di file testuale utilizzato per rappresentare informazioni con struttura tabellare. Le righe delle tabelle corrispondono a righe nel file di testo CSV e i valori delle celle sono divisi da un carattere separatore, che, come indica il nome stesso, dovrebbe essere la virgola. Il CSV non è uno standard vero e proprio ma la sua modalità d'uso è descritta nell'RFC 4180. Nel rilascio di dati secondo il formato CSV, per agevolare i riutilizzatori, si raccomanda di dichiarare almeno 1) il separatore di campo utilizzato (per es., virgola, punto e virgola); 2) se è stato usato un carattere per delimitare i campi di testo.

Raccomandazioni sul formato CSV[58]

- **Dichiara il separatore di campo utilizzato**
  Come indicato nel nome stesso del formato, il separatore dei campi è generalmente la virgola. Tuttavia, possono essere utilizzati altri segni; il documento europeo suggerisce di utilizzare il punto e virgola perché la virgola è utilizzata spesso all'interno dei valori, per esempio in caso di numeri decimali. In ogni caso, è bene rendere esplicito al fruitore dei dati qual è il separatore utilizzato.

[58] Tratte dal documento "data.europa.eu – Data quality guidelines", indicato nel box "Risorse utili", a cui si rimanda per ulteriori approfondimenti.

- **Utilizza un file per tabella**

  Ogni file CSV deve contenere solo una tabella. Se la tabella da pubblicare è composta da più fogli, è necessario creare un file CSV per ogni foglio. Nella tabella ogni riga deve contenere una sola osservazione, così come ogni colonna una sola variabile (cd. regole di *Tidy Data*).

- **Evita gli spazi bianchi e informazioni aggiuntive nel file**

  È importante assicurarsi che il file contenga solo i dati che appartengono alla tabella effettiva, come le intestazioni di colonna e i valori delle voci presenti nella tabella stessa. Nel file CSV, quindi, non devono essere presenti titolo della tabella, righe vuote o eventuali informazioni aggiuntive che aiutino l'utente a capire meglio i dati (queste ultime, che sono utilissime, vanno inserite nei metadati). Il file, inoltre, deve contenere una sola riga di intestazione.

- **Inserisci le intestazioni di colonna**

  Le intestazioni di colonna devono essere auto-esplicative ed essere incluse nella prima riga del file CSV. Senza le intestazioni, è difficile per gli utenti interpretare il significato dei dati. Le intestazioni dovrebbero seguire le etichette dei concetti definiti nelle ontologie di riferimento, se disponibili, pubblicate nel Catalogo nazionale della semantica dei dati.

- **Assicurati che tutte le righe abbiano lo stesso numero di colonne**

  Ogni riga deve avere lo stesso numero di colonne e, quindi, di caratteri separatori. Se in una riga manca un valore, questo di solito viene interpretato come "null". Ciò può comportare un trattamento errato dei dati. Se il CSV contiene righe con un numero diverso di colonne, bisognerebbe controllare se c'è un problema con valori di 'escape' non corretti (ad es. un valore che corrisponde al carattere separatore che in quel caso non va interpretato come tale).

- **Indica le unità in una modalità facilmente elaborabile**

  L'unità di misura di un valore dovrebbe essere indicata nell'intestazione della colonna. Se l'unità cambia da un valore all'altro, allora bisognerebbe considerare una colonna dedicata con un'opportuna intestazione e non inserire l'unità insieme al valore stesso. Per le unità dovrebbero essere utilizzati i codici (URI) derivati da un vocabolario controllato.

Risorse utili

- CSV on the Web, W3C
- data.europa.eu – Data quality guidelines, Publications Office, 2021

2.1.2. JSON (JavaScript Object Notation)

È un formato aperto per la rappresentazione e lo scambio di dati semi-strutturati, leggibile anche dagli utenti e che mantiene, rispetto a formati simili come l'XML, una sintassi poco prolissa. Questo aspetto ne fa un formato flessibile e compatto. Esso nasce dalla rappresentazione di strutture dati semplici nel linguaggio di programmazione JavaScript, ma mantiene indipendenza rispetto ai linguaggi di programmazione.

Raccomandazioni sul formato JSON[59]

- **Utilizza tipi di dati adeguati**

  JSON consente l'utilizzo dei seguenti tipi di dati:

  - Valore nullo (assenza di un valore), rappresentato dalla parola chiave 'null';
  - Valori booleani, vero o falso;
  - Stringhe, dove la mascheratura dei singoli caratteri funziona allo stesso modo del file CSV;
  - Numeri e sequenze semplici delle cifre 0–9, eventualmente con un segno e/o punto decimale;
  - Elenchi, detti anche array, racchiusi tra parentesi quadre, in cui i singoli elementi sono separati da virgole. Gli elenchi possono anche essere vuoti;
  - Oggetti, racchiusi tra parentesi graffe e contenenti un numero qualsiasi di coppie chiave-valore separate da virgole.

  Per ulteriori elaborazioni è importante utilizzare tipi di dati adeguati.

- **Utilizza le gerarchie per raggruppare i dati**

[59] Tratte dal documento "data.europa.eu – Data quality guidelines", indicato nel box "Risorse utili", a cui si rimanda per ulteriori approfondimenti.

Invece di allegare tutti i campi all'oggetto radice del JSON, i dati dovrebbero essere raggruppati semanticamente. Ciò migliora la leggibilità da parte degli esseri umani e può migliorare le prestazioni durante l'elaborazione del file.

Risorse utili

📖 Generating JSON from Tabular Data on the Web, W3C

📖 data.europa.eu – Data quality guidelines, Publications Office, 2021

2.1.3 XML (eXtensible Markup Language)

È un linguaggio di marcatura standardizzato dal W3C usato per l'annotazione di documenti e per la costruzione di altri linguaggi più specifici per l'annotazione di documenti. XML è basato sull'utilizzo di marcatori (tag) che consentono di strutturare il contenuto informativo da rappresentare. Nell'ambito del Web Semantico è stata definita una specifica serializzazione RDF/XML.

Raccomandazioni sul formato XML[60]

- **Fornisci una dichiarazione XML**
  Ogni file XML dovrebbe avere una dichiarazione XML completa. Questa contiene metadati relativi alla struttura del documento ed è importante affinché le applicazioni elaborino correttamente il file.
- **Fai l' "escaping" dei caratteri speciali**
  Quando vengono utilizzati caratteri speciali nei file XML, è necessario eseguire l' "escape". Ciò garantisce una struttura del file pulita e impedisce alle applicazioni utilizzate per l'elaborazione del file di interpretare erroneamente i dati. L' 'escape' viene eseguito sostituendoli con le entità XML equivalenti.
- **Utilizza nomi significativi per gli identificatori**
  Tutti gli identificatori, siano essi tag o attributi, dovrebbero avere nomi significativi e non dovrebbero auspicabilmente essere usati due volte.

[60] Tratte dal documento "data.europa.eu – Data quality guidelines", indicato nel box "Risorse utili", a cui si rimanda per ulteriori approfondimenti.

- **Utilizza correttamente attributi ed elementi**

  Sebbene non vi sia una direttiva vincolante obbligatoria in merito alla codifica dei dati in elementi o attributi, la prassi è che le informazioni che fanno parte dei dati effettivi debbano essere rappresentate da elementi. I metadati che contengono informazioni aggiuntive dovrebbero invece essere implementati come attributi.

- **Rimuovi i dati specifici del programma**

  XML, come qualsiasi formato aperto, dovrebbe essere sempre indipendente da programmi o strumenti specifici utilizzati per l'elaborazione dei file. Questo permette all'utente di scegliere lo strumento che preferisce per il trattamento dei dati senza doverlo prima bonificare.

2.1.4. Serializzazioni RDF

N-triples

È una serializzazione di RDF in cui ogni tripla è espressa interamente e indipendentemente dalle altre. La concatenazione delle triple di un dataset RDF secondo N-Triples avviene utilizzando il carattere punto (es., <soggetto1> <predicato1> <oggetto1> . <soggetto2> <predicato2> <oggetto2>).

Notation3

Notation3 (o N3) è una serializzazione RDF pensata per essere più compatta rispetto a quella ottenuta utilizzando la sintassi XML. Essa risulta più leggibile da parte degli utenti e possiede delle caratteristiche che esulano dall'uso stretto di RDF (per es., rappresentazione di formule logiche).

Turtle

È una versione semplificata (un sottoinsieme di funzionalità) di N3. Un dataset in Turtle è una rappresentazione testuale di un grafo RDF e, al contrario di RDF/XML, è di più facile lettura e gestione anche manuale.

JSON-LD

È un formato di serializzazione per RDF, standardizzato dal W3C, che fa uso di una sintassi JSON. Viene proposto come formato per Linked Data, mascherando di proposito la sua natura di serializzazione di RDF per ragioni di diffusione del formato. Il gruppo di lavoro che l'ha definito

ha posto come obiettivo, oltre quello di mettere a disposizione un'ulteriore funzionalità al framework RDF, anche quello di avvicinare il mondo dello sviluppo Web e degli utilizzatori dei sistemi di gestione dati NoSQL (in particolare dei document store) al Web Semantico. Da un punto di vista pratico è possibile rilasciare dati RDF utilizzando questo «dialetto» JSON nelle situazioni in cui inizialmente non ci si possa dotare di tecnologie ad-hoc come triple store. Allo stesso tempo, con JSON-LD si fornisce uno strumento standard che consente il collegamento di documenti JSON che per loro natura sono unità di informazione indipendenti.

Raccomandazioni sul formato RDF/xxx[61]

- **Utilizza URI http/https per identificare le risorse**
  Gli ID di una risorsa dovrebbero essere URI HTTP/HTTPS, poiché questi consentono l'accesso diretto alla risorsa in questione. Rendono inoltre le risorse indicizzabili dai motori di ricerca, il che migliora la loro reperibilità.
- **Utilizza 'namespace' (spazi dei nomi) quando possibile**
  Sebbene gli spazi dei nomi non siano necessari per l'elaborazione di RDF, riducono la verbosità e le dimensioni del file.

Risorse utili

- Generating RDF from Tabular Data on the Web, W3C
- data.europa.eu – Data quality guidelines, Publications Office, 2021

2.2 Formati aperti più diffusi per i dati geografici

2.2.1 Shapefile

È il formato standard de-facto per la rappresentazione dei dati dei sistemi informativi geografici (GIS). I dati sono di tipo vettoriale. Lo shapefile è stato creato dalla società privata ESRI che rende comunque pubbliche le sue specifiche. L'apertura delle specifiche ha consentito lo sviluppo di diversi strumenti in grado di gestire e creare tale formato. Seppur impropriamente ci si riferisca a uno shapefile, nella pratica si devono considerare almeno tre file: un .shp contenente le

[61] Tratte dal documento "data.europa.eu – Data quality guidelines", indicato nel box "Risorse utili", a cui si rimanda per ulteriori approfondimenti.

forme geometriche, un .dbf contenente il database degli attributi delle forme geometriche e un file .shx come indice delle forme geometriche. A questi tre si deve anche accompagnare un file .prj che contiene le impostazioni del sistema di riferimento. Si raccomanda comunque di specificare nei metadati la proiezione utilizzata, meglio se fornendo il file .prj stesso. È importante notare che non risulta ancora chiaro se tale formato lo si possa considerare propriamente aperto (e quindi coerente con la definizione introdotta dal **CAD**) di livello 3 secondo il modello per i dati proposto nel presente documento. Tenuto conto dell'ampio uso di tale formato per la rappresentazione dei dati geografici si ritiene opportuno includerlo comunque in questo elenco.

2.2.2. KML

È un formato basato su XML per rappresentare dati geografici. Nato con Google, è diventato poi uno standard OGC. Le specifiche della versione 2.2 presentano una serie di entità XML attraverso cui archiviare le coordinate geografiche che rappresentano punti, linee e poligoni espressi in coordinate WGS84 e altre utili a definire gli stili attraverso cui visualizzare i dati. Eventuali attributi delle geometrie vanno espressi invece attraverso la personalizzazione di alcune entità. Molti strumenti di conversione non si occupano tuttavia di creare questa struttura dati e delegano gli attributi delle geometrie allo stile di visualizzazione. Si consiglia pertanto di distribuire questo dato prestando attenzione o, eventualmente, accompagnando il dataset assieme ad un altro formato aperto per i dati geografici (ad esempio, .shp, .geojson).

2.2.3. GeoJSON

È un formato aperto per la rappresentazione e l'interscambio dei dati territoriali in forma vettoriale, basato su JSON. Ogni dato è codificato come oggetto che può rappresentare una geometria, una caratteristica o una collezione di caratteristiche. A ogni oggetto è associato un insieme di coppie nome/valore (membri). I principali nomi di membri che rappresentano le caratteristiche dei dati geografici sono: «type» che serve a indicare il tipo di geometria (punto, linea, poligono o insieme multi-parte di questi tipi); «coordinates» attraverso cui sono indicate le coordinate dell'oggetto in un dato sistema di riferimento; «bbox» attraverso cui sono indicate le coordinate di un riquadro di delimitazione geografica; «crs» (opzionale) per l'indicazione del sistema di riferimento. Inoltre, è possibile associare all'oggetto specifici attributi, attraverso il membro con nome «properties». Si tratta di un formato molto diffuso e supportato da diversi software, ampiamente utilizzato in ambito di sviluppo web. Nel 2016 è stata pubblicata la relativa RFC 7946 "The GeoJSON Format". La specifica raccomanda di limitare la precisione delle coordinate a 6 decimali, attraverso cui si può specificare qualsiasi posizione sulla terra con una

tolleranza di 10 centimetri. La specifica inoltre richiede che i dati siano memorizzati con un sistema di riferimento di coordinate geografiche WGS 84, in latitudine e longitudine, nello stesso stile dei dati GPS.

2.2.4. GML (Geography Markup Language)

È una grammatica XML che rappresenta un formato di scambio aperto per i dati territoriali. Definita originariamente da OGC e diventata poi lo Standard ISO 19136:2008, essa fornisce la codifica XML (schemi XSD) delle classi concettuali definite in diversi Standard ISO della serie 19100 e di classi aggiuntive quali: geometrie, oggetti topologici, unità di misura, tipi di base, riferimenti temporali, caratteristiche, sistemi di riferimento, copertura.

2.2.5. GeoPackage

È un formato aperto per la rappresentazione di dati geografici e può essere un'alternativa al suddetto formato shapefile. Esso supporta SpatiaLite ovvero un'estensione dello schema del database SQLite. Il principale vantaggio offerto da GeoPackage è quello di rappresentare in un unico file diversi dati geografici, sia di tipo vettoriale che raster, che possono essere gestiti anche tramite apposite interrogazioni SQL. Lo standard è riconosciuto dall'Open Geospatial Consortium.

Risorse utili

- GeoJSON encoding of INSPIRE datasets
- GeoPackage encoding of INSPIRE datasets

2.3 Formati aperti per i documenti

2.3.1 ODF (Open Document Format)

È uno standard dell'OASIS che specifica le caratteristiche di un formato per documenti digitali basato su XML, indipendente dall'applicazione e dalla piattaforma utilizzata. La seguente serie di formati aperti è parte dello standard OASIS ODF:

- ODT (Open Document Text). Standard aperto per documenti testuali. È stato adottato come formato principale per i testi in alcune suite per l'automazione

d'ufficio come OpenOffice.org e LibreOffice; è supportato da altre come Microsoft Office, Google Drive e IBM Lotus.

- ODS (Open Document Spreadsheet). Standard aperto per fogli di calcolo. Come nel caso precedente, è stato adottato come formato principale per i fogli di calcolo in alcune suite per l'automazione d'ufficio come OpenOffice.org e LibreOffice; è supportato da altre come Microsoft Office, Google Drive e IBM Lotus.
- ODP (Open Document Presentation). Standard aperto per documenti di presentazione. È stato adottato come formato principale per i documenti di presentazione in alcune suite per l'automazione d'ufficio come OpenOffice.org e LibreOffice; è supportato da altre come Microsoft Office, Google Drive e IBM Lotus.

2.3.2 PDF

È un formato aperto creato da Adobe per la rappresentazione di documenti contenenti testo e immagini che sia indipendente dalla piattaforma di lettura (applicativo, sistema operativo e hardware). È stato standardizzato dall'ISO (ISO/IEC 32000-1:2008) con una serie di formati differenti, ognuno avente una propria prerogativa (per es., PDF/UA per l'accessibilità, PDF/H per documenti sanitari, PDF/A per l'archiviazione, ecc.). Si noti che rilasciare dati secondo tale formato limita fortemente il riutilizzo dei dati stessi in quanto l'intervento umano richiesto per la loro elaborazione è molto elevato (dati rilasciati in formato PDF con una licenza aperta rappresentano solo il **primo livello del modello dei dati aperti e quindi non sono coerenti con il REQUISITO 2 delle presenti Linee Guida**).

2.3.3 Akoma Ntoso

È un linguaggio basato su XML per la rappresentazione di documenti giuridici. Nel 2017 è diventato una specifica OASIS.

Lo standard Akoma Ntoso 1.0 al primo livello è stato adottato quale standard di riferimento per la rappresentazione elettronica dei provvedimenti normativi con la Circolare n. 2/2019 di AgID recante "*Adozione di standard per la rappresentazione elettronica e l'identificazione univoca del patrimonio informativo di natura giuridica e istituzione del Forum Nazionale per l'informazione giuridica*" (v. box "Risorse utili"), a cui si rimanda per tutto quello non previsto nelle presenti Linee Guida.

Risorse utili – par. 2.3

- Circolare AgID n. 2/2019 "*Adozione di standard per la rappresentazione elettronica e l'identificazione univoca del patrimonio informativo di natura giuridica e istituzione del Forum Nazionale per l'informazione giuridica*"

2.4 Altri formati per i dati di elevato valore

Per le serie di dati di elevato valore, il Regolamento UE dispone che, in generale, si debba utilizzare un formato aperto e leggibile meccanicamente riconosciuto nell'Unione o a livello internazionale, indicazione che può trovare attuazione seguendo il **REQUISITO 2** e quanto riportato innanzi nel presente allegato.

Per alcune categorie tematiche, il predetto Regolamento indica la possibilità di utilizzare anche alcuni formati specifici che sono riportati di seguito.

2.4.1 Formati per dati meteorologici

Oltre che, in generale, un formato aperto e leggibile meccanicamente riconosciuto nell'Unione o a livello internazionale, per i **dati di osservazione misurati dalle stazioni meteorologiche**, il Regolamento indica che è possibile utilizzare i seguenti formati:

- **JSON** per dati orari;
- **BUFR (Binary Universal Form for the Representation of meteorological data)**, formato di dati gestito dall'Organizzazione Meteorologica Mondiale (WMO – World Meteorological Organization)[62];
- **NetCDF (Network Common Data Form)**, insieme di librerie software e formati di dati indipendenti dalla macchina che supportano la creazione, l'accesso e la condivisione di dati scientifici *array-oriented*[63];

[62] https://community.wmo.int/activity-areas/wmo-codes/manual-codes/bufr-edition-3-and-crex-edition-1

[63] https://www.unidata.ucar.edu/software/netcdf/

- **ASCII (American Standard Code for Information Interchange)**, codice per la codifica dei caratteri da utilizzare per lo scambio generale di informazioni tra sistemi di elaborazione e comunicazione[64].

Per i **dati climatici**, possono essere utilizzati i formati **NetCDF** e **JSON**.

Per gli **avvisi meteo** possono essere utilizzati i seguenti formati:

- **CAP (Common Alerting Protocol)**, formato di dati basato su XML per lo scambio di avvisi pubblici ed emergenze tra tecnologie di allerta[65];
- **RSS (Really Simple Syndication)/Atom**, formati di dati basati su XML per distribuire contenuti come elenchi di informazioni conosciuti come "feed"[66].

Per i **dati radar**, oltre al JSON, può essere utilizzato il formato HDF5 (Hierarchical Data Format), progettato per archiviare e organizzare grandi quantità di dati[67].

Per i **dati del modello NWP (Numerical weather prediction)**, oltre al **JSON** e a **NetCDF**, si può utilizzare il formato **GRIB (General Representation of fields In Binary)**, rappresentazione binaria di dati risultanti da un'osservazione o da una simulazione su modello numerico di una proprietà osservabile in un dominio spaziale e temporale su un sistema di riferimento geospaziale o celeste[68].

2.4.2 Formati per dati statistici

Per i dati statistici il Regolamento indica che, oltre a CSV, JSON e qualsiasi altro formato aperto e leggibile meccanicamente, si può utilizzare anche il formato XML con riferimento a **SDMX (Statistical Data and Metadata eXchange)**, uno standard ISO progettato per descrivere dati statistici e relativi metadati, normalizzare il loro scambio e migliorare la loro condivisione tra organizzazioni statistiche e simili[69].

2.4.3 Formati per i dati relativi alle imprese e alla proprietà delle imprese

Oltre all'indicazione di utilizzare qualsiasi formato che sia aperto e leggibile meccanicamente, per dati e documenti che rientrano nel campo di applicazione del Regolamento Delegato (UE)

---

[64] https://datatracker.ietf.org/doc/html/rfc20
[65] http://docs.oasis-open.org/emergency/cap/v1.2/CAP-v1.2-os.pdf
[66] Per RSS v. https://www.rssboard.org/rss-2-0-1, per Atom v. https://datatracker.ietf.org/doc/html/rfc4287
[67] https://www.hdfgroup.org/solutions/hdf5/
[68] https://old.wmo.int/extranet/pages/prog/www/WMOCodes/ManualonCodes.html#Codes
[69] https://sdmx.org/?page_id=5008

2018/81579 della Commissione il Regolamento indica di utilizzare il formato **XHTML (Xtensible HyperText Markup Language)**, un linguaggio di marcatura per creare pagine web che utilizza la sintassi XML[70].

2.5 Formati per immagini e dati raster

2.5.1 PNG (Portable Network Graphics)

È un formato che supporta la compressione dei dati senza perdita di informazioni (*lossless*). Nel 2004 è stato definito come Standard ISO/IEC 15948:2004[71], rivisto e confermato l'ultima volta nel 2021.

2.5.2 JPEG

È un formato che utilizza un metodo per la compressione con perdita di informazioni (*lossy*) per le immagini digitali. È stato definito come Standard ISO/IEC 10918 composto da cinque parti e denominato ufficialmente "*Information technology – Digital compression and coding of continuous-tone still images*"[72].

2.5.3 JPEG 2000

È un sistema di codifica delle immagini che utilizza tecniche di compressione all'avanguardia basate sulla tecnologia wavelet[73]. Esso utilizza compressione sia *lossy* che *lossless*. È definito dalla suite di Standard ISO/IEC 15444-1:2019 "*Information technology — JPEG 2000 image coding system*".

2.5.4 GeoTIFF (Geographic Tagged Image File Format)

È un formato che definisce una serie di tag per descrivere tutte le informazioni geografiche associate alle immagini TIFF che provengono da sistemi satellitari, ortoimmagini, modelli di elevazione digitali o come risultato di analisi geografiche. È diventato uno Standard OGC[74].

2.5.5 SAFE (Standard Archive Format for Europe)

---

[70] https://www.w3.org/TR/2018/SPSD-xhtml-basic-20180327/
[71] https://www.iso.org/standard/29581.html
[72] https://jpeg.org/jpeg/index.html
[73] https://jpeg.org/jpeg2000/
[74] https://www.ogc.org/standards/geotiff

È un formato standard per l'archiviazione e il trasporto dei dati di osservazione della Terra all'interno del sistema di archiviazione dell'Agenzia spaziale europea (ESA), conforme allo standard Open Archival Information System (OAIS)[75].

[75] https://public.ccsds.org/pubs/650x0m2.pdf

Allegato C
Riepilogo di requisiti e raccomandazioni

Di seguito l'elenco dei requisiti e delle raccomandazioni definiti nelle presenti Linee Guida. Per ciascuno di essi, oltre al numero e al testo, sono indicati i destinatari tenuti alla loro osservanza e gli articoli del Decreto di riferimento (solo per i requisiti).

1. Requisiti

n.	Testo requisito	Destinatari	Art. Decreto
1	I soggetti di cui al par. 1.3 DEVONO rendere disponibili i documenti e i dati di cui al par. 1.1 per il loro riutilizzo a fini commerciali e non commerciali secondo quanto indicato nelle presenti Linee Guida.	Pubbliche Amministrazioni Organismi di diritto pubblico Imprese pubbliche Imprese private	Art. 1, cc. 2, 2-ter, 2-quater
2	I dati DEVONO essere resi disponibili in formato aperto e leggibile meccanicamente ad un livello di almeno 3 stelle nella classificazione del modello di cui all'allegato A.	Pubbliche Amministrazioni Organismi di diritto pubblico Imprese pubbliche Imprese private	Art. 6, cc. 1 e 7
3	Nel caso in cui un dato sia disponibile in più formati, almeno uno di essi DEVE essere coerente con il Requisito 2.	Pubbliche Amministrazioni Organismi di diritto pubblico Imprese pubbliche Imprese private	Art. 6, cc. 1 e 7
4	Nel caso in cui, per soddisfare richieste di apertura, rendere disponibili i dati per il riutilizzo comporti attività che vanno al di là della semplice manipolazione che implicherebbero difficoltà sproporzionate, il titolare dei dati non ha l'obbligo di adeguare i documenti o crearne nuovi o fornire estratti di documenti, motivando, attraverso un apposito provvedimento, le difficoltà sproporzionate anche indicando le attività sui dati che eccedono la semplice manipolazione.	Pubbliche Amministrazioni Organismi di diritto pubblico Imprese pubbliche Imprese private	Art. 6, cc. 2 e 7
5	I dati dinamici DEVONO essere messi a disposizione per il riutilizzo attraverso API coerenti con il Requisito 27.	Pubbliche Amministrazioni Organismi di diritto pubblico Imprese pubbliche Imprese private	Art. 6, cc. 5 e 7
6	I dati dinamici DEVONO essere resi disponibili per il riutilizzo immediatamente dopo la raccolta. Se, per motivi di capacità finanziarie o tecniche, ciò non sia possibile, allora i dati dinamici POSSONO essere resi disponibili per il riutilizzo entro un termine e/o con temporanee restrizioni tecniche, tali, però, da non pregiudicare indebitamente lo sfruttamento del loro potenziale economico e sociale. Il termine e le restrizioni tecniche di cui sopra DEVONO essere definiti e motivati con apposito provvedimento del titolare del trattamento dei dati.	Pubbliche Amministrazioni Organismi di diritto pubblico Imprese pubbliche Imprese private	Art. 6, cc. 5, 6 e 7
7	Le serie di dati di elevato valore DEVONO essere messe a disposizione per il riutilizzo attraverso API coerenti con il Requisito 27 e, se del caso, attraverso download in blocco.	Pubbliche Amministrazioni Organismi di diritto pubblico Imprese pubbliche	Art. 6, c. 8 Art. 12-bis, c. 1
8	Le serie di dati di elevato valore, individuate dal Regolamento (UE) di esecuzione n. 2023/138,	Pubbliche Amministrazioni Organismi di diritto pubblico	Art. 12-bis, c. 1

	DEVONO essere messe a disposizione gratuitamente ad eccezione di quelle: • in possesso delle imprese pubbliche qualora ciò determini una distorsione della concorrenza nei pertinenti mercati, se così previsto dagli atti di esecuzione di cui sopra; • in possesso di biblioteche, comprese le biblioteche universitarie, musei e archivi; • che, in possesso degli enti pubblici, generano utili per coprire una parte sostanziale dei costi inerenti allo svolgimento dei compiti di servizio pubblico e la loro messa a disposizione gratuita avrebbe un impatto sostanziale sul bilancio di tali enti. In tal caso la possibile esenzione dall'obbligo di mettere a disposizione gratuitamente le serie di dati di valore elevato deve durare per un periodo non superiore ai due anni dall'entrata in vigore del Regolamento (UE) di esecuzione n. 2023/138.	Imprese pubbliche	
9	I dati della ricerca DEVONO essere resi disponibili gratuitamente per il riutilizzo per fini commerciali o non commerciali, nel rispetto della disciplina sulla protezione dei dati personali.	Pubbliche Amministrazioni Organismi di diritto pubblico Imprese pubbliche	Art. 7, c. 9-bis
10	Per rendere i dati reperibili È RICHIESTO: • documentare i dati con metadati dettagliati; • assegnare a dati e metadati un identificatore univoco e persistente (per es. DOI); • rendere disponibili online i metadati attraverso una risorsa consultabile, come per esempio un catalogo o un repository, seguendo il Requisito 29 e il Requisito 30; • specificare l'identificatore del dato nei metadati seguendo le indicazioni presenti nei documenti di cui al Requisito 15 e al Requisito 16.	Pubbliche Amministrazioni Organismi di diritto pubblico Imprese pubbliche	Art. 9-bis, c. 3
11	Per rendere i dati accessibili È RICHIESTO: • consentire l'accesso a dati e metadati a partire dall'identificatore univoco e persistente assegnato; • utilizzare protocolli standardizzati e aperti (per es. https); • rendere sempre disponibili i metadati anche quando i dati non sono accessibili (o perché sono applicati meccanismi di autenticazione e autorizzazione all'accesso o perché non più disponibili).	Pubbliche Amministrazioni Organismi di diritto pubblico Imprese pubbliche	Art. 9-bis, c. 3
12	Per rendere i dati interoperabili È RICHIESTO: • fornire i dati in formato aperto secondo i requisiti definiti nelle presenti Linee Guida; • utilizzare gli standard pertinenti per i metadati; • utilizzare, ove possibile, vocabolari controllati, parole chiave, thesauri e ontologie; • inserire riferimenti qualificati ad altri dati o metadati.	Pubbliche Amministrazioni Organismi di diritto pubblico Imprese pubbliche	Art. 9-bis, c. 3
13	Per rendere i dati riusabili È RICHIESTO: • rendere disponibili dati accurati e ben descritti con molti attributi pertinenti; • assegnare ai dati una licenza d'uso chiara e accessibile secondo il Requisito 20 e il Requisito 21; • rendere chiaro come, perché, quando e da chi i dati sono stati creati e processati;	Pubbliche Amministrazioni Organismi di diritto pubblico Imprese pubbliche	Art. 9-bis, c. 3

	• seguire i pertinenti standard di dominio per dati e metadati.		
14	Nel caso di dati territoriali, DEVONO essere applicate le regole tecniche adottate nell'ambito del framework definito in applicazione della Direttiva 2007/2/CE e relativo decreto di recepimento, D. Lgs. n. 32/2010, oltre che di tutte le altre norme ad essi collegate.	Pubbliche Amministrazioni Organismi di diritto pubblico Imprese pubbliche Imprese private	Art. 1, c. 2-quinquies Art. 6, c. 9
15	I dati, con esclusione di quelli territoriali, resi disponibili per il riutilizzo DEVONO essere documentati attraverso metadati conformi al profilo DCAT-AP_IT definito con la guida operativa per i cataloghi dati.	Pubbliche Amministrazioni Organismi di diritto pubblico Imprese pubbliche Imprese private	Art. 6, c. 1 Art. 9, c. 1
16	I dati territoriali resi disponibili per il riutilizzo DEVONO essere documentati esclusivamente attraverso metadati conformi alle "Linee Guida recanti regole tecniche per la definizione e l'aggiornamento del contenuto del Repertorio Nazionale dei Dati Territoriali" e le relative guide operative.	Pubbliche Amministrazioni Organismi di diritto pubblico Imprese pubbliche Imprese private	Art. 6, c. 1 Art. 9, c. 1
17	Le pubbliche amministrazioni e gli organismi di diritto pubblico, al fine di facilitare la conservazione dei documenti disponibili per il riutilizzo DEVONO seguire le indicazioni fornite nelle "Linee Guida sulla formazione, gestione e conservazione dei documenti informatici".	Pubbliche Amministrazioni Organismi di diritto pubblico	Art. 9, c. 3
18	In caso di richieste di riutilizzo di dati DEVE essere seguita la procedura di esame ed evasione di dette richieste definita nelle presenti Linee Guida, salvo che il titolare dei dati non abbia individuato una propria specifica analoga procedura. Sono tenute comunque a definire termini e modalità di riutilizzo dei dati le imprese pubbliche, gli istituti di istruzione, le organizzazioni che svolgono attività di ricerca, le organizzazioni che finanziano la ricerca, il DIS, l'AISE e l'AISI. In ogni caso, DEVONO essere rispettati i termini temporali previsti dal Decreto.	Pubbliche Amministrazioni Organismi di diritto pubblico Imprese pubbliche	Art. 5
19	I titolari dei dati disponibili per il riutilizzo NON DEVONO: • utilizzare licenze che non consentano opere derivate o uso commerciale; • utilizzare licenze di tipo proprietario.	Pubbliche Amministrazioni Organismi di diritto pubblico Imprese pubbliche Imprese private	Art. 1, c. 2 Art. 8, cc. 1 e 2
20	Per i nuovi dati aperti nativi, salvo quanto precisato nel Requisito 21, DEVE essere applicata la licenza CC-BY nell'ultima versione disponibile (al momento della stesura delle presenti Linee Guida, la 4.0), presupponendo altresì l'attribuzione automatica di tale licenza nel caso di applicazione del principio "open data by default", di cui all'articolo 52 del CAD.	Pubbliche Amministrazioni Organismi di diritto pubblico Imprese pubbliche Imprese private	Art. 8, cc. 1 e 2
21	L'adozione, qualora possibile e/o previsto, di una licenza diversa dalla CC-BY 4.0 o CC0 DEVE essere formalmente motivata, anche alla luce dei principi espressi dalla Direttiva, salvo che sia stata adottata una licenza altrettanto compatibile come la CDLA 2.0 permissive ovvero qualsiasi altra licenza aperta equivalente o meno restrittiva, che consenta il riutilizzo salvo obbligo di attribuzione, dando credito al concedente. Per le serie di dati di elevato valore vale quanto indicato nel Regolamento (UE) di esecuzione n. 2023/138.	Pubbliche Amministrazioni Organismi di diritto pubblico Imprese pubbliche Imprese private	Art. 8, cc. 1 e 2
22	I dati DEVONO essere resi disponibili per il riutilizzo gratuitamente, salvo eventuale applicazione dei costi marginali effettivamente sostenuti per la riproduzione, la messa a disposizione e la divulgazione dei dati, nonché per	Pubbliche Amministrazioni Organismi di diritto pubblico	Art. 7, c. 1

	l'anonimizzazione di dati personali o per le misure adottate per proteggere le informazioni commerciali a carattere riservato.		
23	Nel caso in cui sia richiesto il pagamento di un corrispettivo, il totale delle entrate provenienti dalla fornitura e dall'autorizzazione al riutilizzo dei documenti in un esercizio contabile NON PUÒ superare i costi marginali del servizio reso (comprendenti i costi di raccolta, produzione, riproduzione, diffusione, archiviazione dei dati, conservazione e gestione dei diritti e, ove applicabile, di anonimizzazione dei dati personali e delle misure adottate per proteggere le informazioni commerciali a carattere riservato), maggiorati di un utile ragionevole sugli investimenti.	Pubbliche Amministrazioni Organismi di diritto pubblico	Art. 7, c. 3-bis
24	L'importo totale delle tariffe DEVE essere calcolato in base a parametri oggettivi, trasparenti e verificabili ed è determinato secondo il criterio del costo marginale del servizio con decreti dei Ministri competenti, di concerto con il Ministro dell'economia e delle finanze sentita l'Agenzia per l'Italia digitale.	Pubbliche Amministrazioni Organismi di diritto pubblico	Art. 7, c. 4
25	Nel caso di enti territoriali ed enti e organismi pubblici diversi da quelli indicati, gli importi delle tariffe, calcolati sulla base dei criteri indicati innanzi, e le relative modalità di versamento sono determinati con disposizioni o atti deliberativi dell'ente titolare, sentita l'Agenzia per l'Italia Digitale.	Pubbliche Amministrazioni Organismi di diritto pubblico	Art. 7, c. 9
26	Le condizioni applicabili al riutilizzo dei dati e l'effettivo ammontare delle tariffe applicate, compresa la base di calcolo utilizzata per tali tariffe e gli elementi presi in considerazione nel calcolo di tali tariffe, DEVONO essere pubblicati sui siti istituzionali di pubbliche amministrazioni, organismi di diritto pubblico e imprese pubbliche competenti, previa comunicazione ad AgID.	Pubbliche Amministrazioni Organismi di diritto pubblico Imprese pubbliche	Art. 7, c. 9-ter
27	Le API sviluppate per rendere disponibili i dati per il riutilizzo DEVONO essere conformi alle "Linee Guida sull'interoperabilità tecnica delle Pubbliche Amministrazioni" e le "Linee Guida Tecnologie e standard per la sicurezza dell'interoperabilità tramite API dei sistemi informatici", adottate con la Determinazione di AgID n. 547/2021.	Pubbliche Amministrazioni Organismi di diritto pubblico Imprese pubbliche Imprese private	Art. 6, cc. 5 e 8
28	Nel caso di dati territoriali, il Requisito 27 è attuato attraverso l'implementazione dei servizi di rete di cui all'art. 11 della Direttiva 2007/2/CE, del Regolamento (CE) n. 976/2009 e delle relative Linee Guida tecniche.	Pubbliche Amministrazioni Organismi di diritto pubblico Imprese pubbliche Imprese private	Art. 6, c. 9
29	Le amministrazioni DEVONO inserire e mantenere aggiornati nel portale dati.gov.it, attraverso le modalità di alimentazione previste dal catalogo, i metadati dei dati, ad esclusione di quelli territoriali.	Pubbliche Amministrazioni Organismi di diritto pubblico Imprese pubbliche Imprese private	Art. 9, cc. 1 e 2
30	I dati territoriali DEVONO essere documentati esclusivamente presso il Repertorio Nazionale dei Dati Territoriali (RNDT) che, in maniera automatizzata, si occupa dell'allineamento con il portale nazionale dei dati aperti dati.gov.it.	Pubbliche Amministrazioni Organismi di diritto pubblico Imprese pubbliche Imprese private	Art. 9, cc. 1 e 2
31	I destinatari delle presenti Linee Guida DEVONO pubblicare e aggiornare annualmente sui propri siti istituzionali gli elenchi delle categorie di dati detenuti ai fini del riutilizzo.	Pubbliche Amministrazioni Organismi di diritto pubblico Imprese pubbliche Imprese private	Art. 9, cc. 1

Tabella 8 - Requisiti, destinatari e relativi riferimenti normativi

2. Raccomandazioni

n.	Testo raccomandazione
1	SI RACCOMANDA un percorso graduale verso la produzione nativa di Linked Open Data – LOD (livello cinque stelle nella classificazione del modello di cui all'allegato A).
2	Ove possibile, opportuno o necessario, SI RACCOMANDA di rendere disponibili i dati dinamici anche attraverso download in blocco.
3	Ove possibile, SI RACCOMANDA di seguire e applicare i principi FAIR per tutte le tipologie di dati, non solo per quelli della ricerca.
4	SI RACCOMANDA di demandare al Responsabile per la transizione digitale (RTD) il compito di costituire un gruppo di lavoro dedicato al processo di apertura dei dati e all'implementazione delle presenti Linee Guida all'interno dell'organizzazione dell'Ente.
5	SI RACCOMANDA di costituire, all'interno dell'organizzazione dell'Ente, un apposito gruppo di lavoro dedicato al processo di apertura dei dati anche per l'applicazione delle presenti Linee Guida, prevedendo, ove possibile, le strutture e le figure adatte e necessarie a tale scopo, tenendo in considerazione i referenti dei singoli domini (come le infrastrutture di dati territoriali) e prevedendo altresì il necessario coinvolgimento del Responsabile per la protezione dei dati laddove siano coinvolti dati personali.
6	SI RACCOMANDA che le pubbliche amministrazioni statali e regionali che detengono archivi contenenti dati di interesse territoriale – fermo restando il rispetto della disciplina in materia di protezione dei dati personali – prevedano una diffusione dei dati pubblici in forma disaggregata per favorire l'ottimale gestione delle funzioni di raccolta ed elaborazione dati delle Province e Città Metropolitane a supporto di tutti gli enti locali del loro territorio.
7	SI RACCOMANDA di definire un percorso di apertura dei dati da inserire nel Piano Triennale ICT della singola Amministrazione, la cui definizione può rientrare nei compiti da assegnare al RTD. Tale percorso potrà essere basato su una scala di priorità nell'apertura tenendo in considerazione gli obblighi derivanti dall'applicazione del Decreto per alcune specifiche tipologie di dati, fermo restando il rispetto della normativa unionale e nazionale in materia di protezione dei dati personali. Potranno essere considerate anche specifiche regole di politiche dei dati aperti (Data policy), purché non in contrasto con le strategie e le politiche nazionali.
8	SI RACCOMANDA di garantire, per tutti i dati in generale e per quelli resi disponibili per il riutilizzo, in particolare, il rispetto almeno delle quattro caratteristiche di qualità dei dati (ovvero accuratezza, coerenza, completezza e attualità), fra le 15 previste dallo Standard ISO/IEC 25012 come da indicazioni della Determinazione Commissariale n. 68/2013 di AgID. Per la misura delle suddette caratteristiche, fare riferimento allo Standard ISO/IEC 25024.
9	SI RACCOMANDA di restringere le condizioni di cui alla licenza apposta ai dati alla sola attribuzione, fatta eccezione per le regole in materia di riutilizzo dei dati personali (cfr. par. 4.1 e par. 5.1.2).
10	SI RACCOMANDA di limitare l'uso di licenze con condizioni ulteriori rispetto alla sola attribuzione solo ai casi strettamente necessari.
11	SI RACCOMANDA di limitare l'uso della clausola di "condivisione" ("share-alike" - SA) solo ai casi in cui sia motivatamente necessaria ovvero previa verifica di impossibilità di rilascio con licenza CC BY 4.0, ad esempio, in ragione dell'uso non altrimenti gestibile di una fonte già rilasciata con licenza SA.
12	SI RACCOMANDA di non utilizzare le licenze Creative Commons precedenti alla 4.0, in cui i diritti sui generis non erano citati/previsti (2.5), o erano richiamati come meramente rinunciati (3.0).
13	SI RACCOMANDA di evitare quelle licenze che – per quanto ben impostate – presentano forti caratteristiche di localizzazione, anch'esse potenzialmente costituenti elementi di ambiguità in caso di riutilizzo e mashup (come la IODL).
14	SI RACCOMANDA ai titolari che hanno già pubblicato set di dati con licenze diverse da quelle sopra richiamate, incluse versioni della CC-BY precedente alla 4.0, di valutare il rinnovo della licenza, adeguandola alle indicazioni suddette, individuando nel caso le ragioni eventualmente impedienti tale aggiornamento.
15	Ove possibile, SI RACCOMANDA di utilizzare API conformi al Requisito 27 per rendere disponibili per il riutilizzo tutte le tipologie di dati, non solo quelli dinamici e/o di elevato valore.
16	Al fine di poter documentare i dati statistici, compresi quelli identificati come dati ad elevato valore, nel portale nazionale dati.gov.it secondo il relativo profilo di metadati basato sulla specifica europea DCAT-AP, SI RACCOMANDA di fare riferimento all'estensione StatDCAT-AP.
17	SI RACCOMANDA di non creare tanti portali diversi per singole iniziative ma, ove possibile, di raccordarle per facilitare il reperimento e il riutilizzo dei dati da parte degli utenti finali.

Allegato D
Elenco analitico delle risorse utili

Di seguito l'elenco delle risorse utili citate nelle presenti Linee Guida. In corrispondenza di ciascuna risorsa vengono indicate le sezioni del documento (capitoli e paragrafi) dove tali risorse sono indicate come riferimento.

Risorsa	Rif. LG
Architectures and Standards for Spatial Data Infrastructures and Digital Government	Allegato B
Catalogo nazionale della semantica dei dati	5.1.4, 5.1.5
CDLA permissive compatibility	6.1
CDLA-Permissive-2.0 Compatibility with Other Licenses	6.1
Circolare AgID n. 2/2019 "*Adozione di standard per la rappresentazione elettronica e l'identificazione univoca del patrimonio informativo di natura giuridica e istituzione del Forum Nazionale per l'informazione giuridica*"	Allegato B
Circolare del Ministro per la pubblica amministrazione n. 3 del 1° ottobre 2018 - Responsabile per la transizione digitale - art. 17 decreto legislativo 7 marzo 2005, n. 82 "Codice dell'amministrazione digitale"	5.1.1
Compatibility of Creative Commons Licenses	6.1
Compatible Licenses, Creative Commons	6.1
Comunicazione della Commissione Europea 2014/C 240/01 - Orientamenti sulle licenze standard raccomandate, i dataset e la tariffazione del riutilizzo dei documenti	6.2, 6.4
Cool URIs for the Semantic Web	7.1.3
Creative Commons Licenses Compatibility Wizard	6.1
CSV on the Web	Allegato B
data.europa.eu – Data quality guidelines	4.1, Allegato B
European Catalogue for ICT standards	Allegato B
Evaluation of Application Programming Interfaces for INSPIRE	7.1.1
FAIR principles	4.4
Generating JSON from Tabular Data on the Web	Allegato B
Generating RDF from Tabular Data on the Web	Allegato B
GeoJSON encoding of INSPIRE datasets	Allegato B
GeoPackage encoding of INSPIRE datasets	Allegato B
Gestione licenze – data.europa.eu	6.1
Guida nazionale per l'implementazione della specifica GeoDCAT-AP	7.2
Guida operativa per i cataloghi dati	4.6
Guida operativa per la compilazione dei metadati RNDT	4.5, 4.6
Guida per la redazione format del Piano Triennale per le pubbliche amministrazioni	5.1.2
Guide tecniche INSPIRE per i servizi di rete	7.1.1
How FAIR are your data? Checklist	4.4
How to make your data FAIR – Guides for Researchers	4.4
INSPIRE Data Specifications	4.5
INSPIRE Technical Specifications for the implementation of Network Services	4.5
Italian Open Data License	6.1
Joinup Licensing Assistant	6.1
Linee guida in materia di trattamento di dati personali contenuti anche in atti e documenti amministrativi effettuato da soggetti pubblici per finalità di pubblicazione e diffusione sul web	1, 4.1, 5.1.2
Linee Guida recanti indicazioni operative ai fini della definizione delle esclusioni e dei limiti all'accesso civico di cui all'art. 5 co. 2 del D. Lgs. 33/2013	1, 5.1.2
Linee Guida per l'interoperabilità semantica attraverso i Linked Open Data	5.1.4
Methodology for data validation 2.0	5.1.6
OGC API – Features as an INSPIRE download service	7.1.1
OGC SensorThings API as an INSPIRE download service	4.2, 7.1.1

Open Data Goldbook for Data Managers and Data Holders - Practical guidebook for organizations wanting to publish Open Data	4.1
OpenAPI Checker	7.1.1
Parere n. 03/2013 sul principio di limitazione della finalità	5.1.2
Piano nazionale di digitalizzazione del patrimonio culturale – Linee Guida per l'acquisizione, la circolazione e il riutilizzo delle riproduzioni dei beni culturali in ambiente digitale	1, 6.2, 6.4
Piano nazionale di digitalizzazione del patrimonio culturale 2022-2023 e relative Linee Guida	6.2, 6.4
Piano Nazionale Infrastrutture di Ricerca (PNIR) 2021 – 2027	4.4
Programma nazionale per la ricerca (PNR) 2021-2027	4.4
Piano nazionale per la scienza aperta (PNSA) 2021-2027	4.4
Regolamento recante i livelli minimi di sicurezza, capacità elaborativa, risparmio energetico e affidabilità delle infrastrutture digitali per la PA e le caratteristiche di qualità, sicurezza, performance e scalabilità, portabilità dei servizi cloud per per la pubblica amministrazione, le modalità di migrazione, nonché le modalità di qualificazione dei servizi cloud per la pubblica amministrazione	7.1.2
Regolamento sui criteri per la fornitura dei servizi di conservazione dei documenti informatici	5.1.5
Regole tecniche nazionali per i dati territoriali	4.5
Sistema di Registri INSPIRE Italia	5.1.4
StatDCAT Application Profile for data portals in Europe	7.2
Strategia Cloud Italia	7.1.1, 7.1.2
Study on persistent URIs, with identification of best practices and recommendations on the topic for the MSs and the EC	7.1.3
The FAIR data principles	4.4
Technical Guidance for INSPIRE Spatial Data Services	4.5
Webinar "Real-time Data"	4.2
Wiki/cc license compatibility	6.1

Tabella 10 - Elenco delle risorse utili